An Investigation of Instructional Effects on Student Growth in Mathematics with Repeated Measures Using Computerized-Adaptive Test

In the Session of

Technical Issues in Development and Implementation of Computerized-Adaptive Test in K-12 Assessments

By

Liru Zhang
Delaware Department of Education

Shudong Wang
Northwest Evaluation Association

Paper Presented at the NCME Annual Conference
April 13-16, 2012, Vancouver, Canada

Introduction

Fueled by the federal *NCLB Act* of 2001 and state educational policy, there has been intense interest in student academic growth. More challenges are presented by the Race to the Top agenda, "(student) growth may be measured by a variety of approaches, but any approach used must be statistically rigorous and based on student achievement data, and may also include other measures of student learning in order to increase the construct validity and generalizability of the information" (2009, pp.37812-37813). Therefore, student learning gains can be used to evaluate the effectiveness of instructional strategies and individual teachers.

The potential of using computerized adaptive test (CAT) in determining student growth have been described in recent reports, such as efficient testing and improved precision (Betebenner and Linn, 2010; Ballou, 2008; Yen, 1986). CAT has begun to be employed in high-stakes K-12 assessment programs. The current study explores the patterns of academic growth with repeated measures in reading and mathematics from a statewide computerized adaptive assessment. Test scores were collected from three mandated administrations at grades 3, 8, and 10 during the 2010-2011 school year. The Multilevel Linear Models are utilized for analyses and estimated growth patterns are comparisons.

Methods of Study

<u>Purpose of Study</u>

This study is an exploration of the effects of instructional time (time-span) on student growth in mathematics and the relationships between instructional time and student level characteristics (e.g., gender and race). Three waves of data from a computerized adaptive test are used to estimate the growth patterns and change rate for students at three grade levels. The empirical growth is examined based on a vertical scaling to help determine suitable models to estimate student growth. The study is an application of the Multilevel Linear Modeling with a fixed occasion design for repeated measures. Model-fit and associated assumptions are evaluated for appropriate use of multilevel growth modeling. Measurement issues in determining student growth in the context of CAT are discussed.

<u>Assessment Instrument</u>

The Delaware Comprehensive Assessment System (DCAS) is an online adaptive test in mathematics. To measure academic growth, each student was required to take DCAS three times throughout the school year. Table 1 displays the calendar for the 2010-2011 statewide assessment. Each test window lasted seven to thirteen weeks. Schools are responsible for scheduling student assessments due to technology facilities and classroom instruction. Students might accomplish a test on multiple instances.

The DCAS mathematics is designed to measure the Delaware Prioritized Content Standards in the item formats of multiple-choice (MC) and machine-scored constructed-response (MSCR) questions that have varying score scales, from 0-2 to 0-4. In mathematics, students answer questions in the content categories of Numeric Reasoning, Algebraic Reasoning, Geometric Reasoning, and Quantitative Reasoning. The three waves of measures were given at the beginning of the school year, mid-year, and near the end of the school year, along with classroom instruction.

The computer algorithm follows two primary criteria, matching the test specifications and matching estimated student ability, for item selection in the adaptive test. Due to the blended nature of the assessment, the accountability scores that are based on 40 to 50 on-grade items serve as the primary indicator for high-stakes accountability system. Student achievement is reported on the vertical scale, approximately from 300 to 1200, for grades 2-10, as well as in four performance levels, *Well Below Standard*, *Below Standard*, *Meets Standard*, and *Advanced*.

Data Source

The three waves of data used in the current analysis were collected from the 2010-2011 DCAS administration at grades 3, 8, and 10. Students who received a valid score on all three measurements are included and their gain scores between measures are calculated. The vertical scale was designed with the mean of 800 and the standard deviation of 100, thus, the gain scores of over 200 in the absolute value are treated as outliers. Each grade sample contains over 8,000 to 9,000 students.

Methods for Analysis

The initial performance level (IPL) is defined by student test scores in the first measure; time-span or time-lag in the unit of week is defined as the duration of instructional time between measures. The analyses include two stages this study:

(1) Empirical Growth: In order to help determine the suitable individual growth model, empirical growth patterns are examined by descriptive statistics for scale scores and scatter plots for the entire grade samples and by gender, race, social-economic status (SES), and the initial performance level (IPL). It is important in repeated measures to ensure that use the same measures on the same scale and apply the same criterion for evaluation. Scaling is a common practice in educational measurement and testing. Concurrent calibration was applied in the 2010 field test to put all items in the pool on the same scale. It is assumed that each administration represents the measurement of the same characteristics across occasions for the comparability of test scores across repeated measures. The score scale itself does not carry any normative-referenced or criterion-referenced information. Students' scale scores or gain scores from multiple measures could be compared with their peers who had the same initial status from a norm-references perspective or compared against the performance standards to determine student growth from a criterion-referenced perspective.

(2) Multilevel Linear Modeling: For decades, the measurement of change scores at the individual level has been long perplexed researchers and educators (Cronbach & Furby, 1970; Rogosa, Brandt, & Zimowski, 1982; Traub, 1994; Meredith & Horn, 2001). In recent years, several distinct statistical techniques are available for the analysis of panel data. Multilevel Models can be used to model change over time in a variable of interest. An overall change function is fitted to the whole sample and the parameters can be allowed to vary. In the current study, individual students might be assumed to show linear growth, for instance from fall to spring, however, the exact intercept and slope could be different across individuals. The main

emphasis of multilevel models is the explanation of variability between students in the parameters that describe their growth curves.

This study uses the Two-Level Individual Growth Models to estimate student growth trajectories in mathematics. The behavior of a Level-1 outcome is examined as a function of both Level-1 and Level-2 predictors. The dependent variables are scale scores in mathematics from three waves of data as a fixed linear effect. The independent variables include Time-span, and three student demographic categories, Gender, Race, and Social-Economic Status (SES).

Model A is the Unconditional Means Model with repeated measures (within individuals) and Model B is the Unconditional Linear Growth Model with a level-1 predictor at student level (Time-span) to estimate growth between students. Model C is an Individual Linear Growth Model with the level-1 predictor (Time-span) and a level-2 predicator, such as to estimate growth by gender overtime. Model D is a Modified Linear Random Coefficient Model with a random intercept and a fixed slope. Model E, a Fully Linear Random Coefficient Model, with both intercept and slope random to estimate growth. The two types of models are displayed below using Gender and Race as an example. An outline of the growth models used in this study is presented in Table 2 and all models are listed in Table 11.

Model A – Unconditional Mean Model:

Level-1: $Y_{ij} = \pi_{0i} + \varepsilon_{ij}$

Level-2: $\pi_{0i} = \gamma_{00} + \xi_{0i}$

Model B – Unconditional Growth Model:

Level-1: $Y_{ij} = \pi_{0i} + \pi_{1i}\ \text{Time} + \varepsilon_{ij}$ (This model is the same for the following models)

Level-2: $\pi_{0i} = \gamma_{00} + \xi_{0i}$

$\pi_{1i} = \gamma_{10} + \xi_{1i}$

Model C1: $\pi_{0i} = \gamma_{00} + \gamma_{01}\ \text{gender}_i + \xi_{0i}$

$\pi_{1i} = \gamma_{10} + \gamma_{11}\ \text{gender}_i + \xi_{1i}$

Model D1: $\pi_{0i} = \gamma_{00} + \gamma_{01}\ \text{genderi} + \gamma_{02}\ \text{race}_{1i} + \xi_0 i$

$\pi_{1i} = \gamma_{10} + \gamma_{11}\ \text{genderi} + \gamma_{12}\ \text{race}_{1i} + \xi_{1i}$

Model E1: $\pi_{0i} = \gamma_{00} + \gamma_{01}\ \text{gender}_i + \gamma_{02}\ \text{race}_{1i} + \xi_{0i}$

$\pi_{1i} = \gamma_{10} + \gamma_{11}\ \text{gender}_i + \xi_{1i}$

Where

$$\varepsilon_{ij} \sim N(0, \sigma_\varepsilon^2)\ and\ \begin{bmatrix}\xi_{0i}\\ \xi_{1i}\end{bmatrix} \sim N\left\{\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}\sigma_0^2 & \sigma_{01}\\ \sigma_{10} & \sigma_1^2\end{bmatrix}\right\}$$

π_{0i}: Intercept of true change for student i in the population

π_{1i}: Slope of the true change for student i in the population

σ_ε^2: Level-1 residual variance across all time (grades) of measurement for student i in the population

γ_{00}: Population average in Level-1 intercept, π_{0i}, for student with level-2 predictor (Gender: Female = 0, Male =1; Race: Hispanic/Latino=1, Black=2, White=3; Socioeconomic status (SES): Yes=0, No=1)

$\gamma_{01:}$ Population average difference in Level-1 intercept, π_{0i}, for an i-unit difference in Level 2 predictor

$\gamma_{10:}$ Population average in Level-1 slope, π_{2i}, for students with a Level-2 predictor value of 0

$\gamma_{11:}$ Population average difference in Level-1 slope, π_{2i}, for a 1-unit difference in Level-2 predictor

σ_0^2: Level-2 residual variance in true intercept, π_0i, across students in the population

σ_1^2: Level-2 residual variance in true slope, π_{1i}, across students in the population

σ_{01}^2: Level-2 residual covariance between true intercept π_0i and true slope π_{1i}, across students in the population

Results and Discussion

This study investigates the patterns of student academic growth with repeated measures using a computerized adaptive test (CAT). The empirical growth patterns on the state assessment in mathematics are first examined for modeling and then estimated by the Multilevel Linear Modeling with a fixed-occasion design for time. The scale scores on the three waves of mandated measures within a school year are used for analysis at the elementary (Grade 3), middle (Grade 8), and high-school (Grade 10) levels. The average trajectories of change (or growth) are estimated according to the predictors of gender (Female=0; Male=1), (race (Hispanic/Latino=0; White=1) (African American=0; White=1), and family social economic status (SES) (Yes=0; No=1). Growth patterns are compared between alternative models and issues of evaluate student growth using CAT are discussed.

Empirical Growth Patterns

Descriptive statistics of test scores, gain scores, and percent of students of each performance level are presented, respectively, in Tables 3 and 4 at the state level by grade, and wave of measure. Since the state assessments were administered three times a year, at the beginning of the school year, mid-year, and near the end of the school year, it is assumed that student growth or the change of test scores is sensitive to classroom instruction. Time-span or time-lag from measure to measure (e.g., fall to spring) is collected for each student in the unit of week. It can be seen in Table 5 that the average instructional time lasted 13-15 weeks between any two measures and 28-29 weeks between the first and the third measures. The correlation matrices are shown in Table 6 for test scores and time-span by grade and the scatter plots of mathematics scores in Figure 1 between measures. The overall impression is that the average scale scores rise over three waves of measures, indicating student growth in mathematics throughout the school yea, with variations among individual students. The empirical trajectories shown in Figures 2 by grade suggest a similar linear or approximately linear trend across the three waves of measure at all the three grade levels in mathematics. However, the different slopes may indicate a vary rate of change. Obviously, the 3rd graders demonstrate the largest growth in 2010-2011 among the three grades; while the 8th graders show a higher rate of change than their high-school peers.

The results of repeated measures are summarized in Table 7 by students' initial performance level (IPL) for descriptive statistics and the average empirical trajectories are exhibit in Figures 3a-3c. A linear or nearly linear growth trend is shown across grades and IPLs in mathematics. In grade 3, a similar slope is observed for all students despite their IPL. However, in grades 8 and 10, IPL-4 students demonstrate more growth than their peers, which consequently enlarges the performance gaps from measure to measure between IPL-4 students and their peers.

The average empirical growth is examined by gender (Table 8; Figure 4), race (Table 9; Figure 5); and SES (Table 10; Figure 6). A similar linear trajectory is observed between male and female students in each grade; however, a steeper slope for grade 3 indicates a higher change rate or more growth for elementary students than students in the middle and high school levels. As the slope becomes flatter from grade to grade, a slower growth is assumed in high school than in middle school and slower than elementary school. The empirical growth information by race and SES is summarized in Tables 8, 9, and 10, respectively, for descriptive statistics. Overall, disadvantage students, such as minority and low SES, perform lower than their peers in all three grades in mathematics. A similar growth pattern as the grade sample suggests that the 3rd graders demonstrate more growth overtime than the 8th and 10th graders across comparison groups. Take a close look, it is found that White students significantly outperform Hispanic/Latino and African American students and Hispanic/Latino students slightly, but consistently outperform African American students at all grades. Moreover, the observed differences in slope by racial group, even slightly, indicate varied growth trajectories in mathematics over time.

Estimate Growth Patterns by Multilevel Modeling

(1) Model-Fit: In statistics, deviance is a quality of fit statistic for a model that is often used for statistical hypothesis testing, which is 2 times the log-likelihood ratio of the reduced model compared to the full model. AIC under the name of "an information criterion" is a measure of the relative goodness of fit of an estimated statistical model. AIC offers a relative measure of the information lost when a given model is used to describe reality and the tradeoff between bias and variance in model construction, or loosely speaking that of precision and complexity of the model. Given a data set, several competing models may be ranked according to their AICs, with the one having the lowest AIC being the best. AIC judges a model by how close its fitted values tend to be to the true values, in terms of a certain expected value. But it is important to realize that the AIC value assigned to a model is only meant to rank competing models and which one is the best among the given alternatives. The absolute values of the AIC for different models have no meaning; only relative differences can be ascribed. The Bayesian information criterion (BIC) is closely related to AIC. Unexplained variables increase the value of BIC; hence, lower BIC implies either fewer explanatory variables, better fit, or both.

    The statistics for model-fit are summarized in Table 12 by grade and model, as well as the differences between the corresponding models. A smaller value of Deviance, AIC, and BIC for all growth models indicates a better estimate student growth than the means model in all three grades. When Time is added to the Unconditional Linear Growth Model B, the fit statistics dropped significantly from Model A, suggesting that the estimated growth by the Unconditional Linear Growth Model is closer to the reality than the Unconditional Means

Model. Inclusion of Race1, Race2, and SES to the Individual Linear Growth Mode C, the size of all fit statistics decreased considerably from Model B. For instance, the value of fit statistics is about 1400 less for Model C2 than Model B in grade 3. Gender, on the other hand, does not really help improving model-fit. Compared the fit statistics between Models C and D, the smaller size of Deviance, AIC, and BIC suggest that Model D provides a better estimates of student growth in mathematics. More specifically, inclusion of student's family social economic status and racial/ethnic category improves the estimates of student growth by Models D3, D4, and D5. Within each grade, both Model D and Model E provide a better model fit than the simple models. However, compare the values of Deviance, AIC, and BIC between the two complex models, the picture becomes blurred as the lower absolute values vary among the three criteria. Perhaps, we could assume that both Model D and E function similarly well due to the trivial differences or we could say that Model E is the best one for grade 3; while Model D is the best one for grades 8 and 10 because of the majority of lower absolute values belong to.

(2) Estimate Growth Patterns by Model A: The results of fitting series of multilevel growth models are summarized in Table 13 by model and grade. The Unconditional Means Model (Model A) "can be viewed as a one-way random effects ANOVA model" (Singer, 1998). With Model A, we estimate the fixed effect, the initial status, and the random effect. Both within-student and between student variances are found statistically significant for all three grades. The intra-class correction of .6267 tells us that 63% of the total variance occurs between students in grade 3. Model A provides a baseline against which to compare more complex models (Singer, 1998). Two alternative models, Model A-Sample 1 and Model A-Sample 2, only include selected racial groups to facilitate the comparison to the complex models.

(3) Estimate Growth Patterns by Model B: Model B is an Unconditional Linear Growth Model, a baseline model for change over time. Inclusion of Tim-span the level-1 model allows us to examine the effect of instruction on the outcome of student growth in mathematics scores. In grade 3, the fixed effects are 618.28 ($p<.0001$) for γ_{00}, representing the estimated average the initial status when Time=0 and 39.004 for γ_{10} ($p<.0001$) representing the estimated average rate of change. All the variance components are found statistically significant. The level-1 variance (874.31; $p<.0001$) indicates that there is still unexplained within-student residual variance in addition to Time. The level-2 variance components are 4501.99 ($p<.0001$) between students in the initial status, 179.54 ($p<.0001$) in the rate of change, and 104.51 ($p<.0001$), the estimated residual covariance between intercept and slope. The same patterns are observed across grades for all B-models. Compare the effects of Time moving from the Unconditional Means Model to corresponding Unconditional Growth Model, it can be seen that the proportion of reduced level-1 within-student variance associate with linear time is 66-67%, 41-42%, and 24-26% across the three unconditional growth models in grades 3, 8, and 10, respectively.

From the estimates of student growth by Model B, we could say that over two-thirds of the within-student variation in mathematics is explained by linear Time at the elementary school level. However, the variance proportion is dropped by 24-26% from elementary to middle school, and then dropped another 16-17% from middle school to high school. In other words, instructional time has larger effect on student growth in mathematics for students at lower grades than their peers in higher grades. Since significant growth is found in both initial status and rate of change, the trajectories by student characteristics (e.g., gender) should be explored.

(4) Estimate Growth Patterns by Model C: Model C is an Individual Linear Growth Model with a student-level covariate (e.g., Gender) added to the level-2 model. Model C allows us to examine both level-1 and level-2 effects simultaneously, as well as cross-level interaction and to explore if there is differentiate effect of Time by Gender. When Gender is introduced to the level-2 sub-models of Model C1, the fixed effect for γ_{00} is 616.28 ($p<.0001$) in grade 3, representing the estimated average intercept as Gender=0 (female); for γ_{01} is 3.8791 ($p=.0091$), representing the average differentiate intercepts between genders and signifying the significant relationship between the covariate and initial status; and for γ_{10} is 39.1239 ($p<.0001$), representing the average slope when Gender=0 (female). Since the standard error is over three times larger than the parameter estimate (γ_{11}=-.2326; $p=.6487$), there is no significant difference in growth rate, the level 1 slope, in mathematics between male and female. The estimated values are nearly identical to the estimates in the unconditional model (Intercept=618.28; Time=39.003). The coefficient for Gender=3.8791 ($p<.009$) signifies the statistically significant relationship between the covariate and initial status. Comparing variance components from Model B, it is found that inclusion of Gender did not help reduce the variance within students, and for the slope and the covariance between intercept and slope. Actually these variances remain unchanged. The same patterns are observed from the estimates for grade 8, where the average intercept (γ_{00}=791.58; $p<.0001$), the average slope (γ_{10}=18.3934), as well as the variance components are still the same as for Model B. In grade 10, the fixed effect is 824.49 for γ_{00} and 12.1121 for γ_{10}, and the estimate of the coefficient for Gender (γ_{01}=4.3471) are all statistically significant ($p<.0001$). However, as the standard error is 43 times larger than the parameter estimate (γ_{11}=-.00958; $p=.9811$), there is no significant different effect of Time between genders on growth rate in mathematics. Compare with the unconditional model, it is found that the estimates for variance components remain unchanged, indicating that Gender does not help reduce within student variance and improve the fit of intercept and slope.

In Model C2, a student-level covariate, SES is introduced to the both level-2 sub-models. Even though the significant fixed effects in grade 3 for the average intercept (γ_{00}=596.07; $p<.0001$), the average slope (γ_{10}=39.1672; $p<.0001$), and the coefficient for SES (γ_{01}=53.7898; $p<.0001$), the over twice larger standard error (.5183) than the parameter estimate (-.3955; $p=.4454$) indicates that no significant effect of SES on the growth rate in mathematics. Similarly, inclusion of SES did not help reduce the variance components compared with Model B. Actually, the covariance increases slightly from 104.51 for Model B to 109.66 for Model C2. In grade 8, all fixed effects are found statistically significant. The estimate for γ_{01} (37.0965; $p<.0001$) indicates a significant difference in the average initial status between disadvantage students and their peers. With respect to growth rate, the parameter estimate (γ_{11}=4.3149; $p<.0001$) signifies a significant effect of SES on the growth rate between Low-SES and Not Low-SES students. Inclusion of SES does improve the fit of the growth rates as it reduces the variance by 6.4% from 72.625 for Model B to 67.969 for Model C2, as well as reduce the residual covariance by 15%. The same patterns are observed in grade 10. The significant effect of SES suggests a different rate of change between disadvantage students and their peers. Reduced variance (-3.7046) from Model B to Model C1 by 7.9% helps to improve the fit of growth rate. Reduce the residual covariance (-30.53) also helps improve the fit of intercept and slope.

Model C3 introduces Race1, including Hispanic/Latino and White students, into the level-2 sub-models for estimate growth patterns. The significant fixed effects are shown in grade 3 for the average intercept (γ_{00}=593.85; p<.0001), the average slope (γ_{10}=40.4125; p<.0001), and the coefficient for Race1 (γ_{01}=47.0562; p<.0001); however, the parameter estimate (γ_{11}=-.7458; p=.3284) suggests no significant difference in the average change rate in mathematics. Actually the covariance increases slightly (-.6.3383) from 72.4611 for Model B1 to 78.7994 for Model C3. In grade 8, all the fixed effects are found statistically significant. The parameter estimates for γ_{01} (26.7464; p<.0001) and for γ_{11} (4.2995; p<.0001) indicates significant different average initial status and growth rates between Hispanic/Latino and White students. Inclusion of Race1 indeed improves the fit of the growth rates and the initial status as it reduces the variance components by 4-6% from Model B1. Similar patters can be seen in grade 10. The significant effect of Race1 (γ_{11}=2.3466; p=.0007) denotes a different growth rate between Hispanic/Latino and White students in mathematics, which slightly reduces the variance components by 2%. Since our research interest focuses on student growth in mathematics through repeated measures, Model B should be set as the baseline for comparison with complex models.

In Model C4, Race2, including African American and White Students, is introduced to the two level-2 sub-models. All the fixed effects are statistically significant for grades 3, 8, and 10. The parameter estimates suggest a significant average initial status with Race2=0 (African American); a differential initial status (γ_{01}=52.236 for grade 3; γ_{01}=36.1020 for grade 8; γ_{01}=33.3864 for grade 10) between African American and White students; a significant rate of change when Race2=0 (African American); and a differentiate rate of change (γ_{11}=2.3955 for grade 3; γ_{11}=4.4899 for grade 8; γ_{11}=4.0310 for grade 10) between the two groups of students. Comparing variance components from the Unconditional Model B2, Race2 indeed reduces the size of variance by 24% for slope in grade 3, by 14% in grade 8, and by 11% in grade 10. In other words, Race2 accounts for 11-24% of the explainable variance in growth rate. The significance of Race2 helps reduce the covariance by 7% in grade 8 and by 8.1% in grade 10.

From the estimates of student growth by Models C1-C4 in mathematics, it is found that there are significant effects of Time on student growth by SES and Race1, respectively, for grades 8 and 10; and by Race2 for all three grades. These significant effects reduce the variance components and helps improve the fit of estimate initial status and the rate of growth by controlling the student-level covariates. The non-significant effects of SES and Race1 in grade 3 reveal that the achievement gaps in mathematics are trivial in early grade(s) between disadvantaged students (Low-SES and Hispanic/Latino students) and their peers. These gaps become wider in middle and high schools. Since there are significant variations in both initial status and the rate of change, we add two student-level factors to examine student growth patterns.

(5) Estimate Growth Patterns by Model D: In Model D1, both Gender and Race1 are introduced to the two level-2 sub-models to examine the controlled effects on student growth in mathematics. The estimates of parameters in grade 3 for γ_{00} (590.62; p<.0001), for γ_{01} (5.9182; p=.0009), for γ_{02} (47.2553; p<.0001), and for γ_{10} (40.4518; p<.0001) indicate significant effects of Gender in different initial status between male and female when controlling Race1, however, the rate of change is not significant (γ_{11}=-.07209; p=.9116 and $\gamma_{12=}$-.7482; p=.3270). Similar result is observed for Race1. Comparing the variance components, it can be seen that within-student variance, as expected, the same from Model B1 to Model D1 since no level-1 predictor added to the within-student level. Take Gender and Race1 together to reduce the variance for intercept by 9.4% from Model B1, but unchanged for slope. Different patterns are found in grade 8, where the estimate parameters show non-significant fixed effects of Gender (γ_{01}=5.9182; p=.5951) and the rate of change between genders by controlling Race1 (γ_{11}=5.9182; p=.2047), but a significant rate of change (γ_{12}=4.3206; p=.0001), between Hispanic/Latino and White. In other words, by controlling Gender, White students' initial mathematics score is 4.3206 points higher than their peers. Reduced variance components by 2% for intercept, by 4.2% for slope, and by 6.1% for both intercept and slope to improve the fit of models. The value of fit statistics becomes smaller, from 276155 for Model B1 to 171871 for Model D1. The same patters as grade 8 are found in grade 10. Reduced variance components by 3.8% for intercept, 8.1% for slope, and by 2% for both intercept and slope helps improve the model-fit for estimates with smaller fit statistics.

In Model D2, Gender and Race2 are introduced to the two level-2 sub-models to examine the controlled effects on student growth in mathematics. In grade 3, the moderate significant level for γ_{01} (2.9612; p=.00561), non-significant effect of Gender by controlling Race 2 (γ_{11}=.001064; p=.9853), but significant effects of Race2 (γ_{12}=2.3955; p=.0001) indicates that a significant rate of change between African American and White. In other words, by controlling Gender, White students' initial mathematics scores is 2.3955 points higher than African American Students. The same patters are observed in both grades 8 and 10. Reduced variance for intercept (15%) and for covariance (24%) helps improve the model fit for estimates, but the variance for rate of change increases lightly (1.38) in grade 3. Significant reduced variance components in grades 8 (by 7-14.5%) and 10 (by 8-12.4%) indeed help improve the estimates of growth trajectories between-student in mathematics.

In Model D3, both SES and Race1 are introduced to the two level-2 sub-models to examine the controlled effects on student growth in mathematics. In grade 3, the estimates of parameters for γ_{00} (588.91; p<.0001), for γ_{01} (37.8978; p=.0009), for γ_{02} (28.9256; p<.0001), and for γ_{10} (40.5716; p<.0001) indicate significant effects of SES in different initial status between Low-SES and Not Low-SES by controlling Race1; however, the rate of change is not significant (γ_{11}=-1.222; p=.0852) and ($\gamma_{12=}$-.1612; p=.8469). Reduced variance is only found for intercept by 16%, the other two variance components actually increase slightly compared with Model B1. In grade 8, all estimates of parameters show statistically significant, indicating significant effects of SES and Race1, as well as controlled differential change rate between Low-SES students and their peers; and between Hispanic/Latino and White students in mathematics. Significant reduced variance by 11.8% for intercept, 14% for covariance, and 7.2% for slope compared with Model B1. The significant effects of SES (γ_{11}=2.7503;

p<.0001) and the non-significant effects of Race1 in grade 10 ($\gamma_{12=}$1.0612; p=1.495) suggest differential rate of change between Low-SES students and their peers; but indistinguishable rates of change between Hispanic/Latino and White students when controlling SES. In this case, reduced variance by 5.4% helps improve the estimate for slope.

In Model D4, both SES and Race2 are introduced to the two level-2 sub-models to examine the controlled effects on student growth in mathematics. The estimates of parameters in grade 3 are significant for γ_{00} (580.37; p<.0001), for γ_{01} (37.46703; p=.0009), for γ_{02} (37.7274; p<.0001), and for γ_{10} (37.3669; p<.0001) indicate significant effects of SES in different initial status between Low-SES and Not Low-SES by controlling Race2. The significant parameter for Race2 (γ_{12}=2.563; p<.0001) suggests a differential rate of change between African American and White students when controlling the effects of SES. The variance is reduced for initial status and both initial status and rate of change by 21% each, but no help for rate of change compared with Model B2. The growth patterns are found very similar for grades 8 and 10. All the estimates parameters are found statistically significant, indicating the significant effects of SES and Race2, as well as controlled distinguishable rate of change between Low SES students and their peers; and between African American and White students in growth in mathematics. All between-student variance components are dropped with a substantial amount, compared with Model B2, for slope by 10.3% in grade 8 and by 12% in grade 10 when two student-level predictors are added to the level-2 sub-models; which consequently improve the model fit for estimates of student growth.

In Model D5, Gender and SES are introduced to the two level-2 sub-models to examine their controlled effects, respectively, on student growth in mathematics. The estimates of parameters in grade 3 are for γ_{00} (594.23; p<.0001), for γ_{02} (53.7696; p<.0001), for γ_{10} (39.2855; p<.0001) indicates significant effects of SES, and moderate significant effects of Gender ($\gamma_{01=}$3.5852; p=.0096). However, no differential rate of change is found statistically significant for both Gender (γ_{11}=-2305; p=.6517) and SES ($\gamma_{12=}$-.3942; p=.4469). The variance for slope actually increases slightly. A similar pattern is observed in grades 8 and 10. The significant parameters suggest the effects of SES and differential rate of change between Low-SES students and their peers when controlling Gender. Compared with Model B, reduced variance components for intercept (by 11.3-12.8%), slope (6.6-7.9%), and for both (10-14.9) in grades 8 and 10 helps improve the model fit.

Mixed findings based on Model D suggest us to simplify the models to the tentative final Model E.

(6) Model E with both intercept and slope random is used to estimate the differential growth trajectories in mathematics by controlling the effects of a student-level predictor. The review of the results by Model E focuses on the growth trajectories by Gender, Race1, Race2, and SES, respectively. Although there is significant difference in level-1 intercept for Gender (γ_{01}) and Race1 (γ_{02}), the rate of change (γ_{11}) is not significantly different between Male and Female after controlling the effects of Race1, Race2, and SES, separately, in Models E1, E3, and E10 across the three grades. Compared with the corresponding Models D1, D2, and D5, the variance components remain unchanged or nearly identical.

In Models E2, the significant difference in level-1 intercepts are found in grades 3 and 10 for Race1 (γ_{02}) and Gender (γ_{01}), and in grade 8 for Race1 only. After controlling Gender, the fixed effects for Race1 (γ_{12}), are found statistically significant in grades 8 and 10, indicating differential rate of change between Hispanic/Latino and White students. The initial mathematics scores are 4.2995 ($p<.0001$) points higher and 2.3466 ($p<.0001$) points higher for White students than their peers in grade 8 and 10 respectively. Variance components unchanged, suggesting a little is lost by eliminating the main effects of Gender on the rate of change. Compared fit statistics from the corresponding Model D1, the value of fit statistics decreases trivially. Similarly in Model E5, the significant difference in level-1 intercepts are found in all three grades for Race1 (γ_{02}) and SES (γ_{01}). After controlling SES, the fixed effects for Race1 (γ_{12}), are found statistically significant in grades 8 and 10, indicating a different rate of change for Hispanic/Latino and White students. The initial mathematics scores are 3.9906 ($p<.0001$) point higher and 3.0292 ($p<.0001$) points higher in grades 8 and 10, respectively, for White students than their peers. Variance components decrease in a trivial amount from Model D3.

In Model E4, the differences in level-1 intercepts (γ_{02}) and in the rate of change (γ_{12}) for Race 2 are found statistically significant in all three grades after controlling the effects of Gender. In each grade, the initial mathematics score is 2.399 ($p<.0001$), 4.4899 ($p<.0001$), and 4.031 ($p<.0001$) points higher for White students than their peers in grades 3, 8, and 10 respectively. Compared with Model D2, variance components remain unchanged and the fit statistics are nearly identical. The same patterns are observed in Model E7, when controlling the effects of SES, estimates parameters for Race2 are significant. Compared with Model D4, variance components are unchanged and fit statistics remain the same.

Models E5, 8, and 9 examine the effects of SES by controlling Race1, Race2, and Gender, separately. In grade 3, no significant effects are found in the rate of change, suggesting no different growth rate between Low-SES and Not Low-SES student in mathematics. In grades 8 and 10, the statistically significant differences in level-1 intercepts (γ_{02}) and in the rate of change (γ_{11} or γ_{12}) are found for SES after controlling the effects of student-level predictors (Gender, Race1 or Race2). In each grade, the initial mathematics score is higher for Not Low-SES students than their peers in both grades across the three models. Compared with the corresponding Models D3, D4, and D5, all variance components remain nearly the same and without improvement of model fit.

Since all level-2 predictors are dichotomous variables, the estimated growth parameters (intercept and slope) are substituted into all the models. The results are displayed in Table 14 by grade and model.

Summary of Findings

This study is an exploration of the effects of instructional time on student growth in mathematics and the relationships between instructional time and student-level characteristics (e.g., gender and race). Three waves of data from a computerized adaptive test are used to estimate the growth trajectories for students at the elementary, middle, and high-school levels. The intended research question is that if all students have the same growth trajectory if there is significant effect of instructional time. Individual students might be assumed to show linear growth from measure to measure, however, the exact intercept and slope could be different across individuals. Through the application of the Two-Level Individual Growth Models, the variability between students can be explained in estimate of parameters that describe individual students' growth patterns. The summary of the results presented in Table 15 by model and grade facilitates the following findings.

(1) The effects of instructional time is found statistically significant ($p<.0001$) on student growth in mathematics in all grades 3, 8, and 10. The results from Model B suggest that 66% of the within-student variation is explained by linear Time in grade 3. The proportion of variance dropped to 41% and 24% in grades 8 and 10, respectively, indicates instructional time has greater effects on academic growth for students at the elementary level than their peers at middle and high school levels.

(2) The effects of Gender on student growth are found not statistically significant in all three grades across models, indicating indistinguishable growth trajectories between male and female students in mathematics. A few exceptions in grade 8 suggest there might be interaction effects between Gender and other student-level predictors. Moderate significant effects of Gender ($p<.10$) are found by Model E3 after controlling the effects of Race2 and by Model E10 after controlling the effects of SES.

(3) In grade 3, the non-significant effects of SES and Race1 across all models signify indistinguishable growth trajectories between Hispanic/Latino and White students, and between Low-SES students and their peers in mathematics. The results also indicate non-significant achievement gaps in the early grade in mathematics.

(4) In grades 8 and 10, however, student family's social economic status (SES) has significant effects ($p<.0001$) on student growth in mathematics. The significant difference in the initial status and the rate of changes signify the distinct growth trajectories between disadvantaged students and their peers. A moderated significant effect of SES (γ_{11}=-1.277; $p<.05$) by Model E6 is found in grade 3 after controlling the effects of Race1. The size of estimated parameter is trivial. Similarly, the effects of Race1 ($p<.0001$) suggest a considerable relationship between instruction time and Race1 in both grades and the differential growth trajectories between Hispanic/Latino and White students in mathematics. It is noted that the effects of Race1 is found non-significant ($p=.1495$) in grade 10 estimated by Model D3.

(5) The effects of Race2 are consistently found significant ($p<.0001$) on student growth at the three grade levels and estimated by all models. Considerable differences in the initial status ($p<.0001$) and the rate of change ($p<.0001$) from estimated parameters support the distinct growth trajectories between African American and White students in mathematics. The results also indicate the significant achievement gaps starting from elementary level through high school. It is important to note that the effects of Race2 are statistically significant after controlling the effects of Gender and SES estimated by Models D2 and D4 with a random intercept and a fixed slope, as well as by Models E4 and E7 with both intercept and slope random.

(6) Providing valid and reliable measures of student academic growth is a critical element in the current educational reform, particularly in the Race to the Top initiative. The findings from the current study signal the concerns regarding how to reliably determine student growth, set reasonable and fair target expectations; and accurately evaluate growth in high-stakes use for individual consequences and teacher evaluation. Statistical models, such as Multilevel Linear Growth Modeling, can be used to estimate effects of other factors, minimize their influences, and identify individuals' growth trajectories (Zhang and Wang, 2010; Hamilton, Stecher & Yuan, 2008; Koretz, 2005). A common drawback of using growth models is missing data. Due to mobility, drop-out, invalid scores or other reasons, some students are excluded from analysis for growth patterns. This limitation could lead to a bias result. The data used for the current analyses are collected from the statewide assessment with three mandated measures in the 2010-2011 school year; the missing data is no more than 4%.

(7) In the repeated measure design, longitudinal data are collected, in which subjects receive a sequence of different treatments (or exposures) and change over time is assessed. It is important to have the same subjects, use the same measure on the same scale, and apply the same criterion for evaluation. Through concurrent calibrations, all test items are put on the same scale in the computerized-adaptive test (CAT). Due to the tailored nature of CAT, students are generally given easy items in the beginning of the school year and harder items in the mid and at the end of the school year, which not only minimize the ceiling and floor effects compared with linear test, but also provide more reliable scores with a smaller conditional standard error of measurement. These advantages of CAT greatly improve the comparability of test scores across occasions in the repeated measure study.

References

Ballou, D. (2008, April). *Test scaling and value-added measurement*. Paper presented at the National Conference on Value-Added Modeling, Madison, WI.

Betebenner, D.W. and Linn, R.L. (2010). Growth in student achievement: Issues of measurement, longitudinal data analysis, and accountability. Exploratory Seminar: *Measurement Challenges within the Race to the Top Agenda*. Center for K-12 Assessment & Performance Management.

Braun, H. (2010). Issues in measuring student growth and conducting productive analyses. Exploratory Seminar: *Measurement Challenges within the Race to the Top Agenda*. Center for K-12 Assessment & Performance Management.

Bray, J. H. and Maxwell, S. E. (1985). *Multivariate analysis of variance*. Series: Quantitative Applications on the Social Science. Sage Publications, Inc.

Cronbach, L .J., & Furby, L. (1970). How should we measure "change"—or should we? Psychological Bulletin, 74, 68-80.

Hamilton, L.S., Stecher, B. M., & Yuan, K. (2008). Standards-based reform in the United States: History, research, and future directions. RAND Corporation.

Koretz, D.M. (2005). Alignment, high stakes, and the inflation of test scores (CSE Report No. 655). Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing.

Meredith, W. & Horn, J. (2001). The role of factorial invariance in modeling growth and change. In L. M. Collins & A. G. Sayer (Eds.), New methods for the analysis of change. Washington, DC: American Psychological Association.

Rogosa, D. R., Brandt, D., & Zimowski, M. (1982). A growth curve approach to the measurement of change. Psychological Bulletin, 92, 726-748.

Singer, J. D. (1998). *Using SAS PROC MIXED to fit Multilevel Models, Hierarchical Models, and Individual Growth Models*. Journal of Educational and Behavior Statistics. Winter 1998, Vol. 24, No. 4, pp.323-355.

Traub, R. E. (1994). Reliability for the social sciences: Theory and applications. Thousand Oaks, CA: Sage Publications.

U.S. Department of Education. (2009). Executive summary, Race to the top assessment program: Notice of public meetings and request for input.
Retrieved from http://www.ed.gov/programs/racetothetop-assessment/executive-summary.pdf.

Yen, W. M. (1986). The choice of scale for educational measurement: An IRT perspective. *Journal of Educational Measurement*, *23*(4), 299–325.

Zhang, L. R. and Wang, S. D. (2010). Using longitudinal data to measure student academic growth. Paper presented at the 2010 NCME Annual Conference. Denver, CO. April 30-May 3, 2010.

Table 1. 2010-2011 Delaware Assessment Calendar

Measure	Start Day	End Day	Duration in Week
2010 Fall	October 11, 2010	December 14, 2010	Over 9 weeks
2011 Winter	January 5, 2011	April 11, 2011	Over 13 weeks
2011 Spring	April 18, 2011	June 3, 2011	7 weeks

Table 2. Outline of Alternative Linear Growth Models Used in This Study

Growth	Model	Level-1 Model	Level-2 Model
Linear	A. Unconditional Means Model	$Y_{ij} = \pi_{0i} + \varepsilon_{ij}$	$\pi_{0i} = \gamma_{00} + \xi_{0i}$
	B. Unconditional Growth Model	$Y_{ij} = \pi_{0i} + \pi_{1i}\text{ Time} + \varepsilon_{ij}$	$\pi_{0i} = \gamma_{00} + \xi_{0i}$ $\pi_{1i} = \gamma_{10} + \xi_{1i}$
	C. Uncontrolled Effect of Gender*	$Y_{ij} = \pi_{0i} + \pi_{1i}\text{ Time} + \varepsilon_{ij}$	$\pi_{0i} = \gamma_{00} + \gamma_{01}\text{ Gender}_i + \xi_{0i}$ $\pi_{1i} = \gamma_{10} + \gamma_{11}\text{ Gender}_i + \xi_{1i}$
	D. Uncontrolled Effect of Gender*	$Y_{ij} = \pi_{0i} + \pi_{1i}\text{ Time} + \varepsilon_{ij}$	$\pi_{0i} = \gamma_{00} + \gamma_{01}\text{ Gender}_i + \gamma_{02}\text{ Race1}_i + \xi_{0i}$ $\pi_{1i} = \gamma_{10} + \gamma_{11}\text{ Gender}_i + \gamma_{12}\text{ Race1}_i + \xi_{1i}$
	E. Uncontrolled Effect of Gender*	$Y_{ij} = \pi_{0i} + \pi_{1i}\text{ Time} + \varepsilon_{ij}$	$\pi_{0i} = \gamma_{00} + \gamma_{01}\text{ Gender}_i + \gamma_{02}\text{ Race1}_i + \xi_{0i}$ $\pi_{1i} = \gamma_{10} + \gamma_{11}\text{ Gender}_i + \xi_{1i}$

Model A: Examine y variance across students without concerning the time-span.
* Take gender as an example

Table 3. Descriptive Statistics for Mathematics Score by Grade

Grade	Measure	Scale Score						
		N	Minimum	Maximum	Mean	SD	Skewness	Kurtosis
3	1	9321	371	916	618.83	71.61	.207	.160
	2	9321	406	948	656.62	77.54	.219	.346
	3	9321	440	919	694.51	78.80	.141	-.097
8	1	9017	596	1137	791.99	57.38	.816	1.810
	2	9017	621	1124	808.86	61.25	.867	1.728
	3	9017	622	1106	828.01	68.60	.634	.534
10	1	8743	618	1142	826.52	51.81	.902	2.371
	2	8743	624	1171	839.14	62.08	.791	1.098
	3	8743	684	1187	850.70	63.97	.914	1.723

Grade	Measure	Gain Scores							
		N	Range	Min.	Max.	Mean	SD	Skewness	Kurtosis
3	M1-M2	9321	364	-166	198	38	40	.053	.854
	M2-M3	9321	382	-185	197	38	45	-.022	.560
	M1-M3	9321	379	-179	200	76	46	-.062	.445
8	M1-M2	9017	352	-153	199	17	34	.051	.644
	M2-M3	9017	340	-147	193	19	36	.114	.849
	M1-M3	9017	360	-164	196	36	38	.164	.669
10	M1-M2	8743	332	-137	195	13	36	-.030	.522
	M2-M3	8743	351	-164	187	12	36	.073	.684
	M1-M3	8743	330	-146	184	24	38	.106	.518

T1-T2: Gain score from Measure-1 to Measure-2

T2-T3: Gain score from Measure-2 to Measure-3

T1-T3: Gain score from Measure-1 to Measure-2

Table 4. Percent of Students in Performance Level by Grade

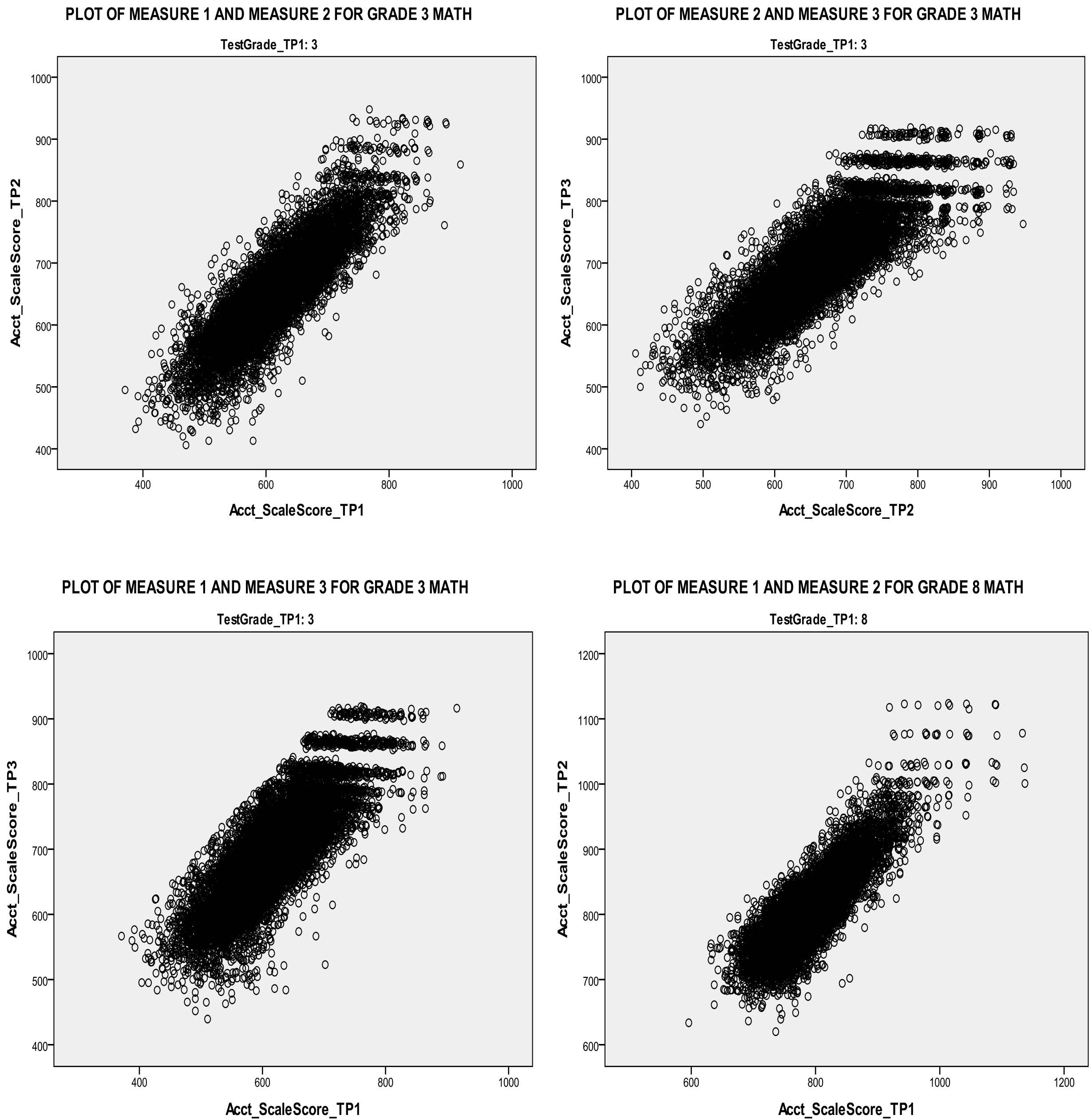
Grade	Mathematics Performance Level	Measure 1 N	Measure 1 %	Measure 2 N	Measure 2 %	Measure 3 N	Measure 3 %
3	*Well below Standard*	3387	36	1864	20	909	10
	Below Standard	3300	35	2963	32	2161	23
	Meet Standard	2262	24	3478	37	4052	43
	Advanced	372	4	1016	11	2199	24
	Total	9321	100	9321	100	9321	100
8	*Well below Standard*	3144	35	2219	25	1659	18
	Below Standard	2286	25	2055	23	1732	19
	Meet Standard	2637	29	3282	36	3083	34
	Advanced	950	11	1461	16	2543	28
	Total	9017	100	9017	100	9017	100
10	*Well below Standard*	2101	24	2029	23	1431	16
	Below Standard	2880	33	2261	26	2063	24
	Meet Standard	3021	35	2995	34	3525	40
	Advanced	741	8	1458	17	1724	20
	Total	8743	100	8743	100	8743	100

Table 5. Descriptive Statistics for Time-Span in Week between Measures

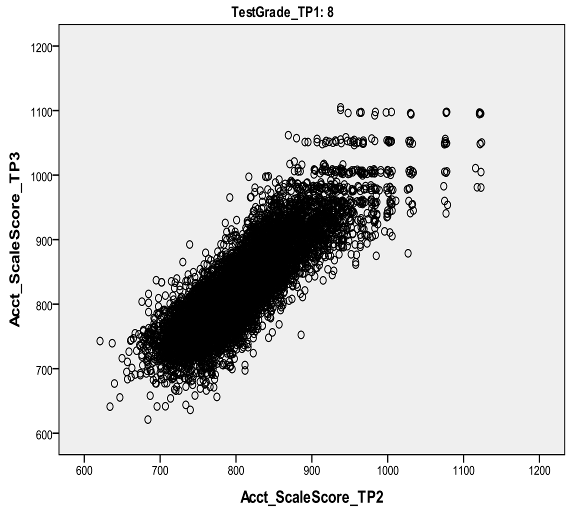
Grade	Time-Span	Mathematics Min.	Mathematics Max.	Mathematics Mean	Mathematics SD
3	Measure 1 to Measure 2	3	23	15	2
	Measure 2 to Measure 3	5	20	14	2
	Measure 1 to Measure 3	18	33	29	2
8	Measure 1 to Measure 2	5	25	15	2
	Measure 2 to Measure 3	4	21	14	2
	Measure 1 to Measure 3	21	33	29	2
10	Measure 1 to Measure 2	4	24	15	2
	Measure 2 to Measure 3	3	22	14	2
	Measure 1 to Measure 3	20	34	28	2

M1-M2: Time-span between two measures, from the beginning day of one measure to the day of finishing the next measure.

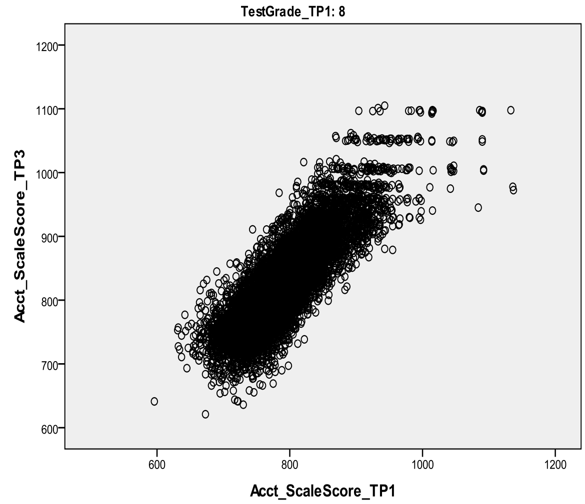
Figure 1. Plot of Scale Scores between Two Measures by Grade



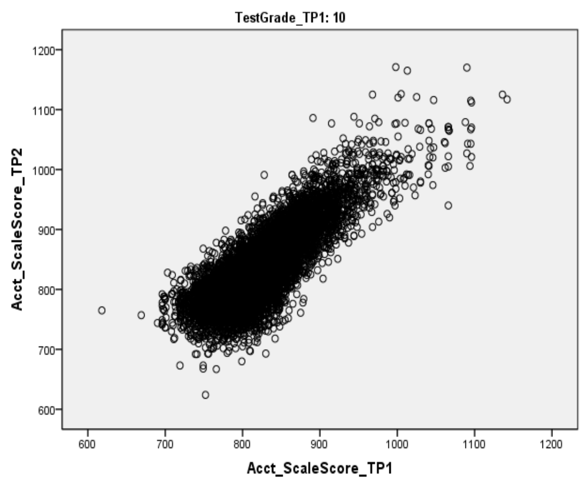
PLOT OF MEASURE 2 AND MEASURE 3 FOR GRADE 8 MATH



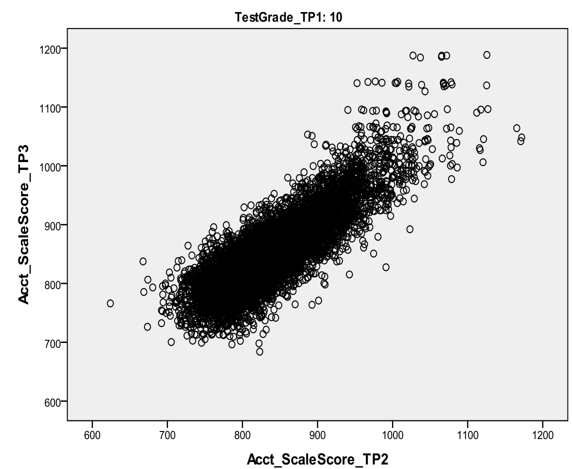
PLOT OF MEASURE 1 AND MEASURE 3 FOR GRADE 8 MATH



PLOT OF MEASURE 1 AND MEASURE 2 FOR GRADE 10 MATH



PLOT OF MEASURE 2 AND MEASURE 3 FOR GRADE 10 MATH



PLOT OF MEASURE 1 AND MEASURE 3 FOR GRADE 10 MATH

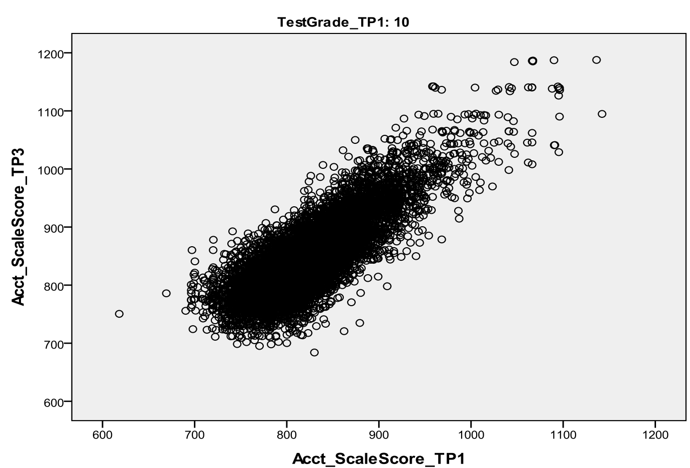


Figure 2. Average Growth Trajectory by Grade

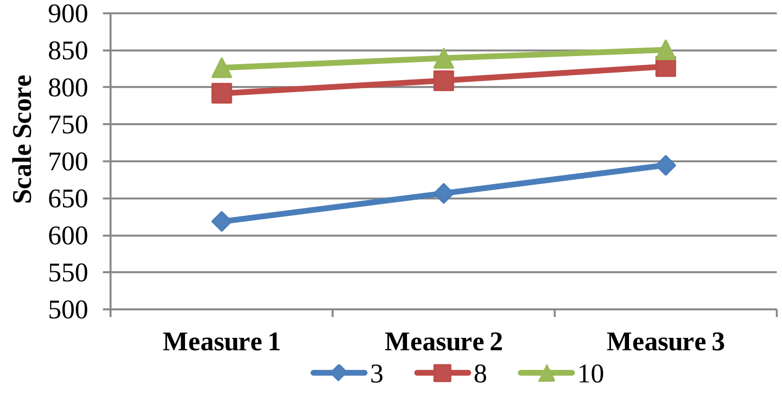


Figure 3a. Growth Trajectyory by IPL - GR 3

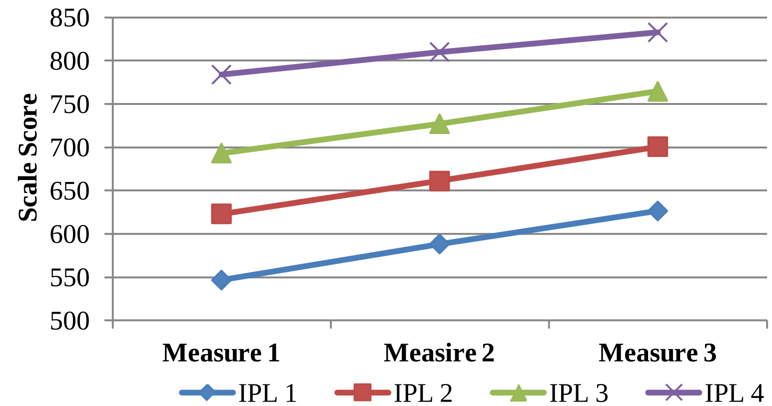


Figure 3b. Growth Trajectory by IPL - GR 8

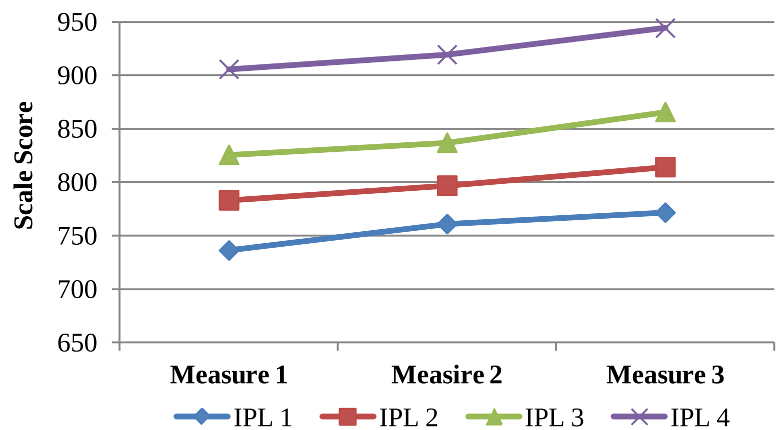


Figure 3c. Growth Trajectory by IPL - GR 10

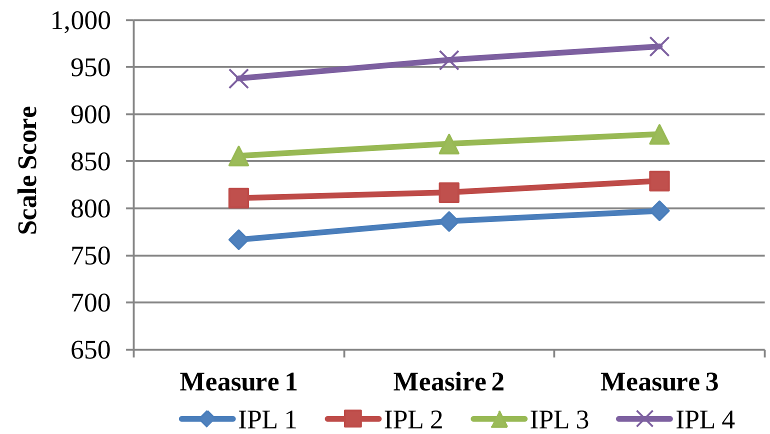


Table 6. Correlation Matrix of Scale Scores

Grade	Measure	Mathematics					
		Score 1	Score 2	Score 3	Time 1	Time 2	Time 3
3	Score 1	1					
	Score 2	.861	1				
	Score 3	.816	.838	1			
	Time-Span 1-2	-.043	.006	-.074	1		
	Time-Span 2-3	.028	-.009	.078	-.691	1	
	Time Span 1-3	-.017	-.005	.010	.338	.446	1
8	Score 1	1					
	Score 2	.842	1				
	Score 3	.834	.850	1			
	Time-Span 1-2	.018	.049	.058	1		
	Time-Span 2-3	-.003	-.029	-.022	-.680	1	
	Time Span 1-3	.018	.020	.042	.322	.475	1
10	Score 1	1					
	Score 2	.820	1				
	Score 3	.809	.840	1			
	Time-Span 1-2	-.015	-.028	-.039	1		
	Time-Span 2-3	.045	.054	.059	-.715	1	
	Time Span 1-3	.036	.029	.020	.471	.280	1

Time-Span 1-2: The weeks between Measire-1 and Measure-2
Time-Span 2-3: The weeks between Measire-2 and Measure-3
Time-Span 1-3: The weeks between Measire-1 and Measure-3

Table 7. Descriptive Statistics for Test Score by Grade and IPL in Mathematics

Grade	IPL	Measure	N	Min.	Max.	Mean	SD	Skewness	Kurtosis
3	1	1	3387	371	592	547	37	-1.044	.844
		2	3387	406	750	588	50	-.318	.235
		3	3387	440	785	626	53	-.057	-.043
	2	1	3300	593	658	623	19	.173	-1.141
		2	3300	490	840	661	41	.055	.715
		3	3300	484	842	701	48	-.068	.491
	3	1	2262	659	749	693	24	.532	-.737
		2	2262	510	934	727	47	.588	1.109
		3	2262	523	917	765	52	.297	.316
	4	1	372	750	916	784	31	1.249	1.418
		2	372	681	948	810	54	.381	-.180
		3	372	678	919	833	53	-.206	-.794
8	1	1	3144	596	766	736	24	-1.094	1.275
		2	3144	621	877	761	34	-.054	.062
		3	3144	622	912	771	37	.032	.374
	2	1	2286	767	799	783	9	-.026	-1.181
		2	2286	650	905	797	32	-.225	.599
		3	2286	670	970	814	39	.062	.484
	3	1	2637	800	861	826	17	.319	-1.012
		2	2637	695	991	837	37	.322	.717
		3	2637	732	1018	866	43	.306	.182
	4	1	950	862	1137	906	42	2.050	5.890
		2	950	780	1124	919	56	.935	1.483
		3	950	817	1106	944	53	.675	.394
10	1	1	2101	618	791	767	20	-1.206	2.155
		2	2101	624	898	786	32	-.086	.593
		3	2101	694	930	797	36	.116	-.040
	2	1	2880	792	829	811	11	-.032	-1.162
		2	2880	680	991	817	36	.028	.296
		3	2880	700	988	829	37	.005	.352
	3	1	3021	830	896	856	18	.432	-.860
		2	3021	693	1086	868	43	.024	.496
		3	3021	684	1049	879	44	.280	.624
	4	1	741	897	1142	938	43	1.830	3.535
		2	741	818	1171	957	52	.799	1.213
		3	741	798	1187	972	62	.706	1.004

IPL: Initial performance level from the first measure in 2010 fall

Table 8. Descriptive Statistics for Test Scores by Grade and Gender

Grade	Gender	Mathematics							
		Measure	N	Min.	Max.	Mean	SD	Skewness	Kurtosis
3	F	1	4525	371	890	617	68	.150	.166
		2	4525	413	948	655	73	.256	.438
		3	4525	440	914	693	75	.126	-.039
	M	1	4796	388	916	621	75	.229	.088
		2	4796	406	934	658	82	.180	.229
		3	4796	452	919	696	82	.142	-.172
8	F	1	4417	632	1137	792	55	.825	2.123
		2	4417	656	1123	809	58	.828	1.809
		3	4417	638	1102	829	66	.601	.505
	M	1	4600	596	1136	792	60	.806	1.541
		2	4600	621	1124	809	64	.893	1.618
		3	4600	622	1106	827	71	.662	.536
10	F	1	4344	618	1094	824	48	.601	1.407
		2	4344	624	1126	837	58	.677	.839
		3	4344	709	1185	849	60	.779	1.334
	M	1	4399	690	1142	829	56	1.049	2.603
		2	4399	673	1171	841	66	.844	1.124
		3	4399	684	1187	853	68	.978	1.794

Table 9. Descriptive Statistics for Test Scores by Grade and Race

Race	Measure	Mathematics						
		N	Min	Max	Mean	SD	Skewness	Kurtosis
Grade 3								
Hispanic/ Latino	1	1365	414	833	594	63	.134	.169
	2	1365	406	928	633	69	.058	.368
	3	1365	481	903	673	70	.103	-.131
African American	1	2964	392	858	590	62	.245	.286
	2	2964	413	930	625	68	.241	.391
	3	2964	440	912	661	72	.257	.156
White	1	4395	371	893	641	69	.023	.318
	2	4395	413	948	680	75	.128	.558
	3	4395	466	919	718	76	.024	.031
Grade 8								
Hispanic/ Latino	1	965	643	955	778	48	.418	.259
	2	965	674	1077	795	53	.818	1.629
	3	965	622	1097	809	59	.569	.821
African American	1	2962	596	964	769	46	.404	.777
	2	2962	634	1006	785	48	.499	.849
	3	2962	642	1053	800	56	.625	.695
White	1	4644	637	1137	806	58	.714	1.454
	2	4644	621	1123	824	62	.741	1.359
	3	4644	645	1106	846	69	.455	.230
Grade 10								
Hispanic/ Latino	1	863	696	974	814	42	.221	.395
	2	863	667	1014	825	51	.511	.626
	3	863	700	1049	836	53	.555	.834
African American	1	2752	618	1142	806	41	.561	2.652
	2	2752	624	1126	813	50	.778	1.549
	3	2752	694	1094	824	51	.734	1.245
White	1	4739	690	1096	837	51	.765	1.701
	2	4739	692	1171	853	62	.674	.868
	3	4739	698	1186	864	63	.768	1.285

Table 10. Descriptive Statistics for Test Score by Grade and SES

Grade	SES	Mathematics							
		Measure	N	Min.	Max.	Mean	SD	Skewness	Kurtosis
3	N	1	3862	410	916	650	70	.114	.268
		2	3862	430	934	689	76	.179	.336
		3	3862	481	919	725	76	.063	-.091
	Y	1	5459	371	837	597	64	.153	.147
		2	5459	406	948	634	71	.162	.480
		3	5459	440	913	673	73	.136	-.055
8	N	1	4321	637	1136	812	60	.836	1.673
		2	4321	621	1123	830	64	.834	1.434
		3	4321	642	1106	852	69	.531	.355
	Y	1	4696	596	1137	774	49	.589	1.541
		2	4696	634	1124	790	51	.700	1.746
		3	4696	622	1102	806	60	.685	.866
10	N	1	4789	696	1142	840	55	.914	2.197
		2	4789	668	1171	856	65	.677	.799
		3	4789	684	1187	868	67	.891	1.497
	Y	1	3954	618	1041	810	42	.451	1.233
		2	3954	624	1125	818	51	.702	1.151
		3	3954	694	1140	830	53	.679	1.260

Figure 4. Growth Trajectory by Gender

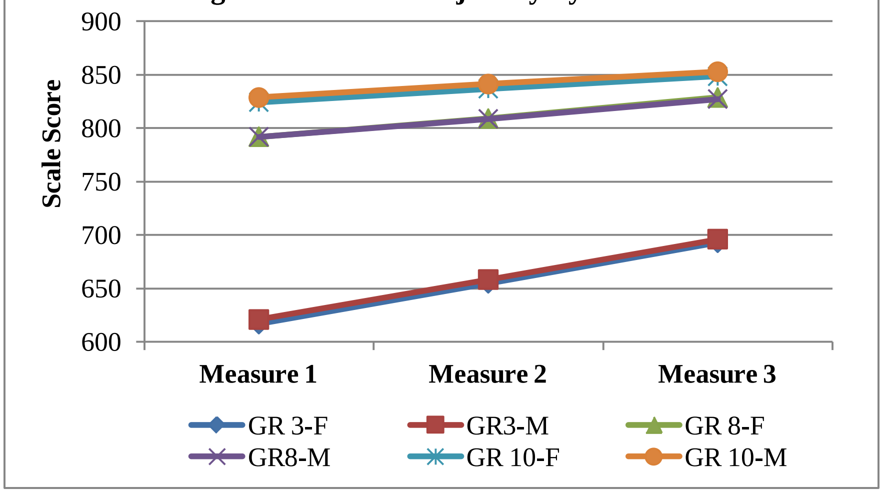


Figure 5. Growth Trajectory by Race

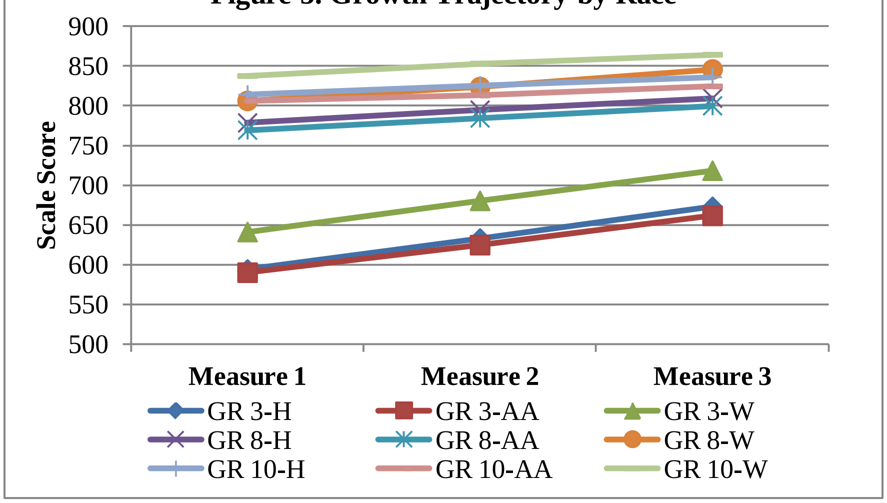


Figure 6. Growth Trajectory by SES

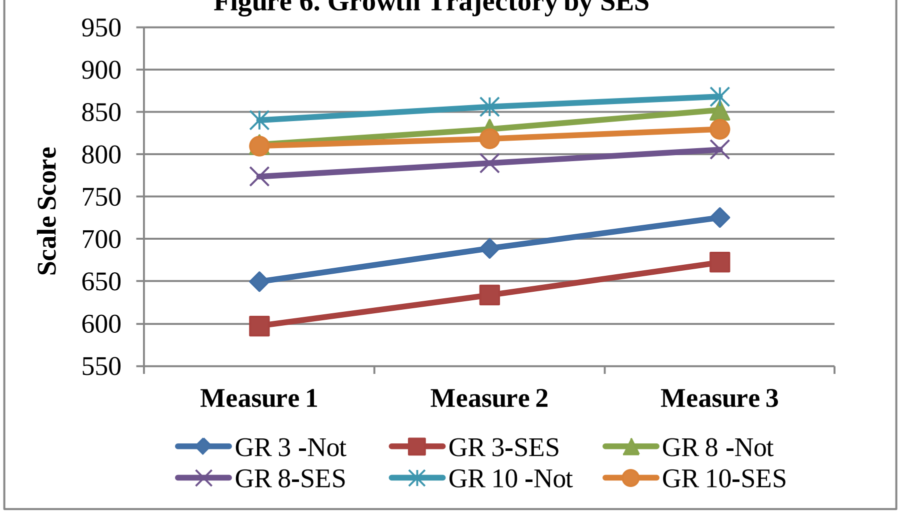


Table 11. List of Models

Model A: Level-1: $Y_{ij} = \pi_{0i} + \varepsilon_{ij}$
Level-2: $\pi_{0i} = \gamma_{00} + \xi_{0i}$

Model B: Level-1: $Y_{ij} = \pi_{0i} + \pi_{1i}$ Time $+ \varepsilon_{ij}$ (The Level-1 sub-model is the same for all the following models)
Level 2: $\pi_{0i} = \gamma_{00} + \xi_{0i}$
$\pi_{1i} = \gamma_{10} + \xi_{1i}$

Model C1: $\pi_{0i} = \gamma_{00} + \gamma_{01}\ \text{gender}_i + \xi_{0i}$
$\pi_{1i} = \gamma_{10} + \gamma_{11}\ \text{gender}_i + \xi_{1i}$
Model C2: $\pi_{0i} = \gamma_{00} + \gamma_{01}\ \text{SES}_i + \xi_{0i}$
$\pi_{1i} = \gamma_{10} + \gamma_{11}\ \text{SES}_i + \xi_{1i}$
Model C3: $\pi_{0i} = \gamma_{00} + \gamma_{01}\ \text{Race1}_i + \xi_{0i}$
$\pi_{1i} = \gamma_{10} + \gamma_{11}\ \text{Race1}_i + \xi_{1i}$
Model C4: $\pi_{0i} = \gamma_{00} + \gamma_{01}\ \text{Race2}_i + \xi_{0i}$
$\pi_{1i} = \gamma_{10} + \gamma_{11}\ \text{Race2}_i + \xi_{1i}$

Model D1: $\pi_{0i} = \gamma_{00} + \gamma_{01}\ \text{gender}_i + \gamma_{02}\ \text{race1}_i + \xi_{0i}$
$\pi_{1i} = \gamma_{10} + \gamma_{11}\ \text{gender}_i + \gamma_{12}\ \text{race1}_i + \xi_{1i}$
Model D2: $\pi_{0i} = \gamma_{00} + \gamma_{01}\ \text{gender}_i + \gamma_{02}\ \text{race2}_i + \xi_{0i}$
$\pi_{1i} = \gamma_{10} + \gamma_{11}\ \text{gender}_i + \gamma_{12}\ \text{race2}_i + \xi_{1i}$
Model D3: $\pi_{0i} = \gamma_{00} + \gamma_{01}\ \text{SES}_i + \gamma_{02}\ \text{race1}_i + \xi_{0i}$
$\pi_{1i} = \gamma_{10} + \gamma_{11}\ \text{SES}_i + \gamma_{12}\ \text{race1}_i + \xi_{1i}$
Model D4: $\pi_{0i} = \gamma_{00} + \gamma_{01}\ \text{SES}_i + \gamma_{02}\ \text{race2}_i + \xi_{0i}$
$\pi_{1i} = \gamma_{10} + \gamma_{11}\ \text{SES}_i + \gamma_{12}\ \text{race2}_i + \xi_{1i}$
Model D5: $\pi_{0i} = \gamma_{00} + \gamma_{01}\ \text{gender}_i + \gamma_{02}\ \text{SES}_i + \xi_{0i}$
$\pi_{1i} = \gamma_{10} + \gamma_{11}\ \text{gender}_i + \gamma_{12}\ \text{SES}_i + \xi_{1i}$

Model E1: $\pi_{0i} = \gamma_{00} + \gamma_{01}\ \text{gender}_i + \gamma_{02}\ \text{race1}_i + \xi_{0i}$
$\pi_{1i} = \gamma_{10} + \gamma_{11}\ \text{gender}_i + \xi_{1i}$
Model E2: $\pi_{0i} = \gamma_{00} + \gamma_{01}\ \text{gender}_i + \gamma_{02}\ \text{race1}_i + \xi_{0i}$
$\pi_{1i} = \gamma_{10} + \gamma_{12}\ \text{race1}_i + \xi_{1i}$
Model E3: $\pi_{0i} = \gamma_{00} + \gamma_{01}\ \text{gender}_i + \gamma_{02}\ \text{race2}_i + \xi_{0i}$
$\pi_{1i} = \gamma_{10} + \gamma_{11}\ \text{gender}_i + \xi_{1i}$
Model E4: $\pi_{0i} = \gamma_{00} + \gamma_{01}\ \text{gender}_i + \gamma_{02}\ \text{race2}_i + \xi_{0i}$
$\pi_{1i} = \gamma_{10} + \gamma_{12}\ \text{race2}_i + \xi_{1i}$
Model E5: $\pi_{0i} = \gamma_{00} + \gamma_{01}\ \text{SES}_i + \gamma_{02}\ \text{race1}_i + \xi_{0i}$
$\pi_{1i} = \gamma_{10} + \gamma_{11}\ \text{SES}_i + \xi_{1i}$
Model E6: $\pi_{0i} = \gamma_{00} + \gamma_{01}\ \text{SES}_i + \gamma_{02}\ \text{race1}_i + \xi_{0i}$
$\pi_{1i} = \gamma_{10} + \gamma_{12}\ \text{race1}_i + \xi_{1i}$
Model E7: $\pi_{0i} = \gamma_{00} + \gamma_{01}\ \text{SES}_i + \gamma_{02}\ \text{race2}_i + \xi_{0i}$
$\pi_{1i} = \gamma_{10} + \gamma_{11}\ \text{SES}_i + \xi_{1i}$
Model E8: $\pi_{0i} = \gamma_{00} + \gamma_{01}\ \text{SES}_i + \gamma_{02}\ \text{race2}_i + \xi_{0i}$
$\pi_{1i} = \gamma_{10} + \gamma_{12}\ \text{race2}_i + \xi_{1i}$

Model E9: $\pi_{0i} = \gamma_{00} + \gamma_{01}\ \text{gender}_i + \gamma_{02}\ \text{SES}_i + \xi_{0i}$

$\pi_{1i} = \gamma_{10} + \gamma_{12}\ \text{SES}_i + \xi_{1i}$

Model E10: $\pi_{0i} = \gamma_{00} + \gamma_{01}\ \text{gender}_i + \gamma_{02}\ \text{SES}_i + \xi_{0i}$

$\pi_{1i} = \gamma_{10} + \gamma_{11}\ \text{gender}_i + \xi_{1i}$

Where $\varepsilon_{ij} \sim N(0, \sigma_\varepsilon^2) and \begin{bmatrix} \xi_{0i} \\ \xi_{1i} \end{bmatrix} \sim N\left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{10} & \sigma_1^2 \end{bmatrix} \right\}$

π_{0i}: Intercept of true change for student i in the population

π_{1i}: Slope of the true change for student i in the population

σ_ε^2: Level-1 residual variance across all time (measures) of measurement for student i in the population

γ_{00}: Population average in Level-1 intercept, π_{0i}, for student with level-2 predictor (Gender: Female = 0, Male =1; Race: Hispanic/Latino=0, Black=0, White=1; Socioeconomic status (SES): Yes=0, No=1)

γ_{01}: Population average difference in Level-1 intercept, π_{0i}, for an i-unit difference in Level 2 predictor

γ_{10}: Population average in Level-1 slope, π_{2i}, for students with a Level-2 predictor value of 0

γ_{11}: Population average difference in Level-1 slope, π_{2i}, for a 1-unit difference in Level-2 predictor

σ_0^2: Level-2 residual variance in true intercept, π_0i, across students in the population

σ_1^2: Level-2 residual variance in true slope, π_{1i}, across students in the population

σ_{01}^2: Level-2 residual covariance between true intercept π_0i and true slope π_{1i}, across students in the population

Table 12. Model-Fit Statistics for Multilevel Linear Models

Model	Grade 3			Difference			Grade 8			Difference			Grade 10			Difference		
	Deviance	AIC	BIC	D	A	B	Deviance	AIC	BIC	D	A	B	Deviance	AIC	BIC	D	A	B
A	320807	320813	320834				284551	284557	284479				271641	271647	271668			
A-S1	198045	198051	198071				177554	177560	177580				174017	174023	174043			
A-S2	253446	253452	253473				239676	239682	239703				232134	232140	232161			
B	303362	303374	303417	17445	17439	17418	276144	276156	276198	8408	8402	-8280	267085	267097	267140	-4556	-4550	-4529
B-S1	187032	187044	187084	11013	11007	10987	172083	172095	172135	5471	5465	-5445	170754	170766	170806	-3263	-3257	-3237
B-S2	239900	239912	239953	13547	13541	13520	232643	232655	232696	7034	7028	-7007	228302	228314	228356	-3831	-3825	-3805
C1	303355	303371	303428	-7	-3	12	276141	276157	276214	-3	1	15	267070	267086	267142	-15	-11	3
C2	301957	301973	302031	-1404	-1400	-1386	274962	274978	275035	1181	1177	-1163	266147	266163	266220	-938	-934	-920
C3	186537	186553	186607	-495	-491	-478	171853	171869	171922	-230	-226	-213	170579	170595	170648	-175	-171	-158
C4	238771	238787	238842	-1129	-1125	-1111	231698	231714	231770	-945	-941	-927	227426	227442	227497	-876	-872	-859
D1	186526	186546	186613	-11	-7	6	171851	171871	171938	-2	2	15	170564	170584	170650	-15	-11	3
D2	238767	238787	238856	-4	0	14	231694	231714	231783	-4	0	14	227414	227434	227504	-12	-8	6
D3	186142	186162	186229	-395	-391	-378	171402	171422	171489	-451	-447	-434	170306	170326	170393	-272	-268	-255
D4	238242	238262	238331	-529	-525	-511	231129	231149	231219	-569	-565	-551	227033	227053	227122	-393	-389	-375
D5	301951	301971	302042	-1404	-1400	-1386	274959	274979	275050	1182	1178	-1164	266137	266157	266227	-934	-930	-915
E1	186527	186545	186605	1	-1	-8	171892	171910	171970	41	39	32	170575	170593	170653	12	10	3
E2	186526	186544	186604	0	-2	-9	171853	171871	171931	2	0	-7	170565	170583	170643	2	-1	-7
E3	238784	238802	238864	17	15	8	231796	231814	231877	102	100	93	227495	227513	227576	81	79	72
E4	238767	238785	238847	0	-2	-9	231698	231716	231778	3	1	-5	227415	227433	227495	0	-2	-9
E5	186145	186163	186223	3	1	-6	171435	171453	171512	32	30	24	170331	170349	170409	25	23	16
E6	186142	186160	186221	0	-2	-9	171418	171436	171496	16	14	7	170309	170327	170386	2	0	-7
E7	238242	238260	238323	0	-2	-9	231183	231201	231264	54	52	45	227074	227092	227154	41	39	32
E8	238258	238276	238338	16	14	7	231176	231194	231257	47	45	38	227072	227090	227152	39	37	30
E9	301951	301969	302033	0	-2	-9	274962	274980	275044	3	1	-6	266137	266155	266218	0	-2	-9
E10	301851	301969	302034	-99	-1	-9	275076	275094	275158	117	115	108	266228	266246	266310	91	89	82

Difference is the discrepancy of absolute values between one model and the previous, corresponding model on the three fit statistics.

Table 13. Results of Fitting Multilevel Growth Models

Model A - Mathematics		Parameter		Grade 3		Grade 8		Grade 10	
Fixed Effect	Initial Status	Intercept	γ_{00}	657.28 (.7397)	<.0001	809.66 (.6220)	<.0001	838.78 (.5963)	<.0001
Variance	Level 1	Within Person		2575.16 (26.4611)	<.0001	975.53 (10.2716)	<.0001	812.67 (8.6893)	<.0001
Component	Level 2	In Initial Status		4323.59 (75.8179)	<.0001	3163.97 (52.0681)	<.0001	2839 (47.1186)	<.0001
Model Fit		Deviance		320806		284451		271641	
Statistics		AIC		320812		284457		271647	
		BIC		320834		284478		271668	
Model A Sample 1 - Math		Parameter		Grade 3		Grade 8		Grade 10	
Fixed Effect	Initial Status	Intercept	γ_{00}	669.61 (.9194)	<.0001	819.82 (.7872)	<.0001	847.53 (.7312)	<.0001
Variance	Level 1	Within Person		2642.85 (34.5656)	<.0001	1028.61 (13.7307)	<.0001	826.67 (11.0429)	<.0001
Component	Level 2	In Initial Status		4060.89 (92.1292)	<.0001	3134.90 (65.8128)	<.0001	2720.88 (56.7267)	<.0001
Model Fit		Deviance		198044		177554		174016	
Statistics		AIC		198050		177560		174022	
		BIC		198070		177580		174042	
Model A Sample 2 - Math		Parameter		Grade 3		Grade 8		Grade 10	
Fixed Effect	Initial Status	Intercept	γ_{00}	658.51 (.8291)	<.0001	809.36 (.6644)	<.0001	837.87 (.6274)	<.0001
Variance	Level 1	Within Person		2568.99 (29.6938)	<.0001	976.25 (11.1917)	<.0001	800.31 (9.2443)	<.0001
Component	Level 2	In Initial Status		4288 (84.6804)	<.0001	3033.51 (54.5844)	<.0001	2683.27 (48.2884)	<.0001
Model Fit		Deviance		253446		239676		232133	
Statistics		AIC		253452		239682		232139	
		BIC		253472		239703		232160	
Model B1 - Math		Parameter		Grade 3		Grade 8		Grade 10	
Fixed Effect	Initial Status	Intercept	γ_{00}	618.28 (.7432)	<.0001	791.62 (.5926)	<.0001	826.68 (.5621)	<.0001
	Rate of Change	Intercept	γ_{10}	39.004 (.2552)	<.0001	18.0387 (.2002)	<.0001	12.1072 (.2019)	<.0001
Variance	Level 1	Within Person		874.31 (12.7052)	<.0001	577.52 (8.5995)	<.0001	619.33 (9.365)	<.0001
Component	Level 2	In Initial Status		4501.99 (76.7432)	<.0001	2686.83 (47.716)	<.0001	2247.69 (42.5143)	<.0001
		In Rate of Change		179.54 (10.9849)	<.0001	72.625 (6.8881)	<.0001	46.7517 (7.1395)	<.0001
		Covariance		104.51 (19.8148)	<.0001	268.59 (12.607)	<.0001	304.65 (11.5994)	<.0001
Model Fit		Deviance		303361		276143		267085	
Statistics		AIC		303373		276155		267097	
		BIC		303416		276198		267139	

Model B2 - Math		Parameter		Grade 3		Grade 8		Grade 10	
Fixed Effect	Initial Status	Intercept	γ_{00}	629.77 (.9307)	<.0001	800.74 (.7522)	<.0001	834.5 (.6892)	<.0001
	Rate of Change	Intercept	γ_{10}	39.8431 (.3243)	<.0001	19.0752 (.2551)	<.0001	13.0293 (.2497)	<.0001
Variance	Level 1	Within Person		881.17 (16.2984)	<.0001	598.88 (11.3056)	<.0001	615.22 (11.6224)	<.0001
Component	Level 2	In Initial Status		4328.98 (94.632)	<.0001	2676.07 (60.676)	<.0001	2149.17 (51.2106)	<.0001
		In Rate of Change		174.21 (13.9899)	<.0001	65.8753 (8.9171)	<.0001	41.6872 (8.7928)	<.0001
		Covariance		72.4611 (24.9413)	<.0001	268.1 (15.4536)	<.0001	300.26 (14.1312)	<.0001
Model Fit		Deviance		187032		172083		170753	
Statistics		AIC		187044		178095		170765	
		BIC		187084		172135		170805	
Model B3 - Math		Parameter		Grade 3		Grade 8		Grade 10	
Fixed Effect	Initial Status	Intercept	γ_{00}	619.81 (.8296)	<.0001	791.29 (.6327)	<.0001	825.96 (.5921)	<.0001
	Rate of Change	Intercept	γ_{10}	38.6995 (.2897)	<.0001	18.0667 (.2176)	<.0001	11.9095 (.2196)	<.0001
Variance	Level 1	Within Person		886.24 (14.4868)	<.0001	578.86 (9.3847)	<.0001	611.8 (9.994)	<.0001
Component	Level 2	In Initial Status		4413.37 (85.0754)	<.0001	1563.28 (49.9934)	<.0001	2117.74 (43.7228)	<.0001
		In Rate of Change		185.09 (12.5665)	<.0001	70.9878 (7.4941)	<.0001	46.6757 (7.625)	<.0001
		Covariance		125.51 (22.3236)	<.0001	265.85 (12.8979)	<.0001	290.85 (12.1903)	<.0001
Model Fit		Deviance		239899		232642		228302	
Statistics		AIC		239911		232654		228314	
		BIC		239953		232696		228355	
Model C1 - Math		Parameter		Grade 3		Grade 8		Grade 10	
Fixed Effect	Initial Status	Intercept	γ_{00}	616.28 (1.0671)	<.0001	791.58 (.8467)	<.0001	824.49 (.7968)	<.0001
		Gender	γ_{01}	3.8791 (1.4865)	<.0091	.0805 (1.1855)	0.9459	4.3471 (1.1233)	,0001
	Rate of change	Intercept	γ_{10}	39.1239 (.3665)	<.0001	18.3934 (.2859)	<.0001	12.1121 (.2864)	<.0001
		Gender	γ_{11}	-.2326 (.5106)	0.6487	-.6954 (.4003)	0.0824	-.00958 (.4037)	0.9811
Variance	Level 1	Within Person		874.31 (12.7052)	<.0001	577.52 (8.5995)	<.0001	619.33 (9.365)	<.0001
Component	Level 2	In Initial Status		4498.23 (76.6891)	<.0001	2686.83 (47.716)	<.0001	2242.97 (42.4441)	<.0001
		In Rate of Change		179.53 (10.9847)	<.0001	72.5042 (6.8867)	<.0001	46.7517 (7.1395)	<.0001
		Covariance		104.73 (19.808)	<.0001	268.61 (12.059)	<.0001	304.66 (7.1395)	<.0001
Model Fit		Deviance		303354		276140		267070	
Statistics		AIC		303370		276156		267086	
		BIC		303428		276213		267142	

Model C2 - Math		Parameter		Grade 3		Grade 8		Grade 10	
Fixed Effect	Initial Status	Intercept	γ_{00}	596.07 (.9025)	<.0001	773.85 (.7755)	<.0001	809.23 (79684)	<.0001
		SES	γ_{01}	53.7898 (1.4046)	<.0001	37.0965 (1.1202)	<.0001	31.8583 (1.0768)	<.0001
	Rate of change	Intercept	γ_{10}	39.1672 (.333)	<.0001	15.9686 (.2756)	<.0001	9.9895 (.2986)	<.0001
		SES	γ_{11}	-.3955 (.5183)	0.4454	4.1393 (.3981)	<.0001	3.8672 (.4034)	<.0001
Variance	Level 1	Within Person		874.31 (12.7052)	<.0001	577.52 (8.5995)	<.0001	619.33 (9.365)	<.0001
Component	Level 2	In Initial Status		3800.64 (66.6636)	<.0001	2343.39 (42.6668)	<.0001	1996.25 (38.7832)	<.0001
		In Rate of Change		179.51 (10.9845)	<.0001	67.969 (6.834)	<.0001	43.0469 (7.0973)	<.0001
		Covariance		109.66 (186167)	<.0001	228.6 (11.428)	<.0001	274.12 (11.1076)	<.0001
Model Fit		Deviance		301957		274962		266147	
Statistics		AIC		301973		274978		266163	
		BIC		302030		275035		266219	
Model C3 - Math		Parameter		Grade 3		Grade 8		Grade 10	
Fixed Effect	Initial Status	Intercept	γ_{00}	593.85 (1.8363)	<.0001	778.6 (1.7846)	<.0001	814.04 (1.7309)	<.0001
		Race 1	γ_{01}	47.0562 (2.1016)	0.0091	26.7464 (1.9611)	0.9459	24.1827 (1.8819)	<.0001
	Rate of change	Intercept	γ_{10}	40.4125 (.6667)	<.0001	15.515 (.6131)	<.0001	11.0440 (.6355)	<.0001
		Race1	γ_{11}	-.7458 (.7630)	0.3284	4.2995 (.6737)	0.0824	2.3466 (.691)	0.0007
Variance	Level 1	Within Person		881.17 (16.2984)	<.0001	598.88 (11.3056)	<.0001	615.22 (11.6224)	<.0001
Component	Level 2	In Initial Status		3929.07 (87.3181)	<.0001	2574.21 (58.7772)	<.0001	2072.98 (49.798	<.0001
		In Rate of Change		174.11 (13.9884)	<.0001	63.2432 (8.8787)	<.0001	40.9698 (8.7826)	<.0001
		Covariance		78.7994 (24.0654)	<.0001	251.73 (15.1978)	<.0001	292.86 (13.9514)	<.0001
Model Fit		Deviance		186537		171853		170578	
Statistics		AIC		186553		171869		170594	
		BIC		186606		171922		170647	

Model C4 - Math		Parameter		Grade 3		Grade 8		Grade 10	
Fixed Effect	Initial Status	Intercept	γ_{00}	588.67 (1.2196)	<.0001	769.24 (.9611)	<.0001	804.84 (.9274)	<.0001
		Race2	γ_{01}	52.2352 (1.5794)	<.0001	36.102 (1.2298)	<.0001	33.3864 (1.166)	<.0001
	Rate of change	Intercept	γ_{10}	37.2712 (.4554)	<.0001	15.3246 (.3465)	<.0001	9.3597 (.3559)	<.0001
		Race2	γ_{11}	2.3955 (.5898)	<.0001	4.4899 (.4434)	<.0001	4.031 (.4475)	<.0001
Variance	Level 1	Within Person		886.24 (14.4868)	<.0001	578.86 (9.3847)	<.0001	611.8 (9.994)	<.0001
Component	Level 2	In Initial Status		3756.53 (74.4627)	<.0001	2253.42 (45.0385)	<.0001	1858.66 (39.5764)	<.0001
		In Rate of Change		183.71 (12.5481)	<.0001	66.1951 (7.4337)	<.0001	42.899 (7.5785)	<.0001
		Covariance		95.3847 (21.0964)	<.0001	227.32 (12.2633)	<.0001	259.57 (11.6396)	<.0001
Model Fit		Deviance		238771		231698		227426	
Statistics		AIC		238787		231714		227442	
		BIC		238842		231769		227497	

Model D1 - Math		Parameter		Grade 3		Grade 8		Grade 10	
Fixed Effect	Initial Status	Intercept	γ_{00}	590.62 (2.0771)	<.0001	778.21 (1.9236)	<.0001	811.61 (1.8541)	<.0001
		Gender	γ_{01}	5.9182 (1.7867)	<.0001	.7872 (1.4809)	<.0001	4.9334 (1.357)	0.0003
		Race1	γ_{02}	47.2553 (2.1005)	<.0001	26.7206 (1.9617)	<.0001	24.1237 (1.8797)	<.0001
	Rate of change	Intercept	γ_{10}	40.4518 (.7548)	<.0001	15.828 (.6607)	<.0001	10.7488 (.6815)	<.0001
		Gender	γ_{11}	-.07209 (.6493)	0.9116	-.6453 (.5087)	0.2047	.5981 (.4988)	0.2305
		Race1	γ_{12}	-.7482 (.7633)	0.327	4.3206 (.6738)	<.0001	2.3394 (.6909)	0.0007
Variance	Level 1	Within Person		881.17 (16.2984)	<.0001	598.88 (11.3056)	<.0001	615.22 (11.6224)	<.0001
Component	Level 2	In Initial Status		3920.34 (87.1585)	<.0001	2574.05 (58.7743)	<.0001	2066.9 (49.6852)	<.0001
		In Rate of Change		174.11 (13.9884)	<.0001	63.1392 (8.8772)	<.0001	40.8804 (8.7813)	<.0001
		Covariance		78.9058 (24.046)	<.0001	251.86 (15.1955)	<.0001	292.12 (13.9365)	<.0001
Model Fit		Deviance		186525		171851		170563	
Statistics		AIC		186545		171871		176583	
		BIC		186612		171937		170650	

Model D2 - Math		Parameter		Grade 3		Grade 8		Grade 10	
Fixed Effect	Initial Status	Intercept	γ_{00}	587.15 (1.4547)	<.0001	769.58 (1.137)	<.0001	802.95 (1.0807)	<.0001
		Gender	γ_{01}	2.9612 (1.5499)	<.0001	-.6796 (1.1997)	<.0001	3.8156 (1.1235)	0.0003
		Race2	γ_{02}	52.2363 (1.5791)	<.0001	36.1097 (1.2298)	<.0001	33.3457 (1.1652)	<.0001
	Rate of change	Intercept	γ_{10}	37.2657 (.5434)	<.0001	15.7316 (.4098)	<.0001	9.4609 (.415)	<.0001
		Gender	γ_{11}	.01064 (.5789)	0.9853	-.8037 (.4325)	0.0631	-.2045 (.4315)	0.6356
		Race1	γ_{12}	2.3955 (.5898)	<.0001	4.499 (.4433)	<.0001	4.0332 (.4475)	0.0007
Variance Component	Level 1	Within Person		886.24 (14.4868)	<.0001	578.86 (9.3487)	<.0001	611.8 (9.994)	<.0001
	Level 2	In Initial Status		3764 (74.4273)	<.0001	2243.31 (45.0366)	<.0001	1855.02 (39.5183)	<.0001
		In Rate of Change		183.71 (12.5481)	<.0001	66.0338 (7.4317)	<.0001	42.8886 (7.5784)	<.0001
		Covariance		95.3768 (12.5481)	<.0001	227.18 (12.2608)	<.0001	259.76 (11.6321)	<.0001
Model Fit Statistics		Deviance		238767		231694		227414	
		AIC		238787		231714		227434	
		BIC		238856		231783		227503	

Model D3 - Math		Parameter		Grade 3		Grade 8		Grade 10	
Fixed Effect	Initial Status	Intercept	γ_{00}	588.91 (1.7933)	<.0001	772.7 (1.7455)	<.0001	8080.46 (1.7302)	<.0001
		SES	γ_{01}	37.8978 (1.8915)	<.0001	32.6208 (1.545)	<.0001	23.2699 (1.4743)	0.0003
		Race1	γ_{02}	28.9256 (2.2253)	<.0001	11.9656 (2.025)	<.0001	13.3069 (1.9661)	<.0001
	Rate of change	Intercept	γ_{10}	40.5716 (.6729)	<.0001	14.9315 (.6198)	<.0001	10.3843 (.6478)	<.0001
		SES	γ_{11}	-1.222 (.7097)	0.0852	3.1285 (.5486)	0.0631	2.7503 (.552)	<.0001
		Race1	γ_{12}	-.1612 (.835)	0.8469	2.8371 (.7191)	<.0001	1.0612 (.7361)	0.1495
Variance Component	Level 1	Within Person		881.17 (16.2984)	<.0001	598.88 (11.3056)	<.0001	615.22 (11.6224)	<.0001
	Level 2	In Initial Status		3629.41 (81.8475)	<.0001	2360.76 (54.8037)	<.0001	1962.93 (47.7603)	<.0001
		In Rate of Change		173.8 (13.9837)	<.0001	61.1538 (8.8483)	<.0001	39.4324 (8.7608)	<.0001
		Covariance		88.462 (23.3713)	<.0001	230.61 (14.7157)	<.0001	279.85 (13.6827)	<.0001
Model Fit Statistics		Deviance		186142		171402		170306	
		AIC		186162		171422		170328	
		BIC		186229		171488		170392	

Model D4 - Math		Parameter		Grade 3		Grade 8		Grade 10	
Fixed Effect	Initial Status	Intercept	γ_{00}	580.37 (1.2325)	<.0001	761.19 (.9976)	<.0001	797.09 (.9982)	<.0001
		SES	γ_{01}	37.4703 (1.6379)	<.0001	27.8949 (1.2434)	<.0001	22.0233 (1.186)	<.0001
		Race2	γ_{02}	37.7274 (1.6514)	<.0001	25.9116 (1.2747)	<.0001	25.5598 (1.2155)	<.0001
	Rate of change	Intercept	γ_{10}	37.3669 (.4762)	<.0001	14.3452 (.3701)	<.0001	8.3154 (.3907)	<.0001
		SES	γ_{11}	-.4326 (.629)	0.4917	3.393 (.4612)	<.0001	2.9678 (.4642)	<.0001
		Race2	γ_{12}	2.563 (.6381)	<.0001	3.2504 (.4729)	<.0001	2.9763 (.4758)	<.0001
Variance	Level 1	Within Person		886.24 (14.4868)	<.0001	578.86 (9.3847)	<.0001	611.8 (9.994)	<.0001
Component	Level 2	In Initial Status		3459.39 (69.6743)	<.0001	2082.68 (42.3311)	<.0001	1754.49 (37.9146)	<.0001
		In Rate of Change		183.67 (12.5476)	<.0001	63.6838 (7.4022)	<.0001	41.0074 (7.5553)	<.0001
		Covariance		98.8147 (20.4899)	<.0001	206.67 (11.9146)	<.0001	245.53 (11.4132)	<.0001
Model Fit		Deviance		238241		231129		227032	
Statistics		AIC		238261		231149		227052	
		BIC		238331		231218		227122	

Model D5 - Math		Parameter		Grade 3		Grade 8		Grade 10	
Fixed Effect	Initial Status	Intercept	γ_{00}	594.23 (1.1478)	<.0001	773.96 (.9604)	<.0001	807.53 (9528)	<.0001
		Gender	γ_{01}	3.58532 (1.3833)	<.0001	-.2303 (1.1195)	0.837	3.4910 (1.0716)	<.0001
		SES	γ_{02}	53.7696 (1.4041)	<.0001	37.0984 (1.1202)	<.0001	31.7634 (1.0765)	<.0001
	Rate of change	Intercept	γ_{10}	39.2855 (.4237)	<.0001	16.3389 (.3412)	<.0001	-.1139 (.4018)	<.0001
		Gender	γ_{11}	-.2305 (.5106)	0.6517	-.7317 (.3978)	0.0659	3.8703 (.4036)	0.7768
		SES	γ_{12}	-.3942 (.5183)	0.4469	4.3254 (.398)	<.0001	619.33 (9.365)	<.0001
Variance	Level 1	Within Person		874.31 (12.7052)	<.0001	577.52 (8.5995)	<.0001	1993.21 (38.7381)	<.0001
Component	Level 2	In Initial Status		3797.42 (66.6176)	<.0001	2343.37 (42.6666)	<.0001	43.0437)	<.0001
		In Rate of Change		179.49 (10.9843)	<.0001	67.8352 (6.8325)	<.0001	274.22 (11.102)	<.0001
		Covariance		109.87 (186105)	<.0001	228.56 911.4262)	<.0001	266136	<.0001
Model Fit		Deviance		301950		274958		266156	
Statistics		AIC		301970		274978		266227	
		BIC		302043		275050			

Model E1 - Math		Parameter		Grade 3		Grade 8		Grade 10	
Fixed Effect	Initial Status	Intercept	γ_{00}	590.96	<.0001	777.75	<.0001	811.52	<.0001
		Gender	γ_{01}	5.9075	0.001	0.7766	0.6	4.9327	0.0003
		Race1	γ_{02}	46.8151	<.0001	27.2876	<.0001	24.2276	<.0001
	Rate of change	Time	γ_{10}	39.8711	<.0001	19.3643	<.0001	12.7206	<.0001
		Gender*Time	γ_{11}	-.05387	0.9339	-.5846	0.2687	0.6127	0.2198
Variance	Level 1	Within Person		881.17	<.0001	598.88	<.0001	615.22	<.0001
Component	Level 2	In Initial Status		3920.37	<.0001	2574.07	<.0001	2066.9	<.0001
		In Rate of Change		174.21	<.0001	65.7895	<.0001	41.5943	<.0001
		Covariance		78.8464	0.001	251.51	<.0001	292.09	<.0001
Model Fit		Deviance		186526		171892		170575	
Statistics		AIC		186544		171910		170593	
		BIC		186604		171969		170653	

Model E2 - Math		Parameter		Grade 3		Grade 8		Grade 10	
Fixed Effect	Initial Status	Intercept	γ_{00}	590.64	<.0001	778.25	<.0001	811.6	<.0001
		Gender	γ_{01}	5.8758	0.0008	0.7025	0.6349	4.9599	0.0003
		Race1	γ_{02}	47.2539	<.0001	26.7234	<.0001	24.1233	<.0001
	Rate of change	Time	γ_{10}	40.4125	<.0001	15.515	<.0001	11.044	<.0001
		Race1*Time	γ_{11}	-.7458	0.3284	4.2995	<.0001	2.3466	0.0007
Variance	Level 1	Within Person		881.17	<.0001	598.88	<.0001	615.22	<.0001
Component	Level 2	In Initial Status		3920.34	<.0001	2574.05	<.0001	2066.9	<.0001
		In Rate of Change		174.11	<.0001	63.2432	<.0001	40.9698	<.0001
		Covariance		78.9045	0.001	251.84	<.0001	292.12	<.0001
Model Fit		Deviance		186525		171852		170565	
Statistics		AIC		186534		171870		170583	
		BIC		186603		171930		170643	

Model E3 - Math		Parameter		Grade 3		Grade 8		Grade 10	
Fixed Effect	Initial Status	Intercept	γ_{00}	586.36	<.0001	769.11	<.0001	802.62	<.0001
		Gender	γ_{01}	2.9617	0.0561	-.6881	0.5663	3.8103	0.0007
		Race2	γ_{02}	53.5653	<.0001	36.8974	<.0001	33.8792	<.0001
	Rate of change	Time	γ_{10}	38.6945	<.0001	18.4544	<.0001	11.992	<.0001
		Gender *Time	γ_{11}	0.00978	0.9865	-.7552	0.0828	-1644	0.7046
Variance	Level 1	Within Person		886.24	<.0001	578.88	<.0001	611.82	<.0001
Component	Level 2	In Initial Status		3754.76	<.0001	2253.37	<.0001	1855.02	<.0001
		In Rate of Change		185.09	<.0001	70.8252	<.0001	46.6522	<.0001
		Covariance		94.6107	<.0001	226.36	<.0001	259.28	<.0001
Model Fit		Deviance		238783		231796		227495	
Statistics		AIC		238801		231814		227513	
		BIC		238863		231876		227575	
Model E4 - Math		Parameter		Grade 3		Grade 8		Grade 10	
Fixed Effect	Initial Status	Intercept	γ_{00}	587.15	<.0001	769.66	<.0001	802.97	<.0001
		Gender	γ_{01}	2.9672	0.0504	-.8203	0.4933	3.7886	0.0007
		Race2	γ_{02}	52.2363	<.0001	36.1113	<.0001	33.346	<.0001
	Rate of change	Time	γ_{10}	37.2712	<.0001	15.3246	<.0001	9.3597	<.0001
		Race2*Time	γ_{11}	2.3955	<.0001	4.4899	<.0001	4.031	<.0001
Variance	Level 1	Within Person		886.24	<.0001	578.86	<.0001	611.8	<.0001
Component	Level 2	In Initial Status		3654.33	<.0001	2253.31	<.0001	1855.02	<.0001
		In Rate of Change		183.71	<.0001	66.1951	<.0001	42.899	<.0001
		Covariance		95.3768	<.0001	227.15	<.0001	259.76	<.0001
Model Fit		Deviance		238767		231897		227414	
Statistics		AIC		238785		231715		227432	
		BIC		238847		231778		227494	

Model E5 - Math		Parameter		Grade 3		Grade 8		Grade 10	
Fixed Effect	Initial Status	Intercept	γ_{00}	589	<.0001	772.59	<.0001	808.41	<.0001
		SES	γ_{01}	37.1975	<.0001	32.2179	<.0001	23.4899	<.0001
		Race1	γ_{02}	29.2606	<.0001	11.6864	<.0001	13.2041	<.0001
	Rate of change	Time	γ_{10}	40.4125	<.0001	15.515	<.0001	11.044	<.0001
		Race1*Time	γ_{11}	-.7458	0.3284	4.2995	<.0001	2.3466	0.0007
Variance	Level 1	Within Person		881.17	<.0001	598.88	<.0001	615.22	<.0001
Component	Level 2	In Initial Status		3629.51	<.0001	2360.82	<.0001	1962.92	<.0001
		In Rate of Change		174.11	<.0001	63.2396	<.0001	40.9684	<.0001
		Covariance		88.2834	0.0002	230.22	<.0001	279.73	<.0001
Model Fit		Deviance		186145		171434		170331	
Statistics		AIC		186163		171452		170349	
		BIC		186223		171512		170408	

Model E6 - Math		Parameter		Grade 3		Grade 8		Grade 10	
Fixed Effect	Initial Status	Intercept	γ_{00}	588.97	<.0001	772.33	<.0001	808.41	<.0001
		SES	γ_{01}	37.9298	<.0001	31.4734	<.0001	23.2476	<.0001
		Race1	γ_{02}	28.8332	<.0001	12.5071	<.0001	13.3918	<.0001
	Rate of change	Time	γ_{10}	40.4761	<.0001	16.8379	<.0001	11.1049	<.0001
		SES*Time	γ_{11}	-1.2777	0.0488	3.9006	<.0001	3.0292	<.0001
Variance	Level 1	Within Person		881.17	<.0001	598.88	<.0001	615.22	<.0001
Component	Level 2	In Initial Status		3629.41	<.0001	2360.79	<.0001	1962.93	<.0001
		In Rate of Change		173.8	<.0001	62.1532	<.0001	39.5611	<.0001
		Covariance		88.458	0.0002	230.42	<.0001	279.84	<.0001
Model Fit		Deviance		186142		171417		170308	
Statistics		AIC		186160		171435		170326	
		BIC		186220		171495		170386	

Model E7 - Math		Parameter		Grade 3		Grade 8		Grade 10	
Fixed Effect	Initial Status	Intercept	γ_{00}	580.42	<.0001	760.96	<.0001	769.91	<.0001
		SES	γ_{01}	37.2327	<.0001	28.6901	<.0001	225397	<.0001
		Race2	γ_{02}	37.8194	<.0001	25.6211	<.0001	253762	<.0001
	Rate of change	Time	γ_{10}	37.2712	<.0001	15.3246	<.0001	9.3597	<.0001
		Race2*Time	γ_{11}	2.3955	<.0001	4.4899	<.0001	4.031	<.0001
Variance	Level 1	Within Person		886.24	<.0001	578.86	<.0001	611.8	<.0001
Component	Level 2	In Initial Status		3459.4	<.0001	2084.8	<.0001	1754.53	<.0001
		In Rate of Change		183.71	<.0001	66.19	<.0001	42.895	<.0001
		Covariance		98.793	<.0001	206.09	<.0001	245.2	<.0001
Model Fit		Deviance		238243		231183		227073	
Statistics		AIC		238260		231201		227091	
		BIC		238322		231263		227153	

Model E8 - Math		Parameter		Grade 3		Grade 8		Grade 10	
Fixed Effect	Initial Status	Intercept	γ_{00}	579.77	<.0001	760.86	<.0001	796.86	<.0001
		SES	γ_{01}	369406	<.0001	27.6301	<.0001	21.848	<.0001
		Race2	γ_{02}	39.1353	<.0001	26.6733	<.0001	26.0777	<.0001
	Rate of change	Time	γ_{10}	38.459	<.0001	15.7521	<.0001	9.6174	<.0001
		SES*Time	γ_{11}	0.5318	0.3609	4.5228	<.0001	3.9748	<.0001
Variance	Level 1	Within Person		886.24	<.0001	578.86	<.0001	611.8	<.0001
Component	Level 2	In Initial Status		3459.8	<.0001	2083.79	<.0001	1754.53	<.0001
		In Rate of Change		185.02	<.0001	65.8727	<.0001	42.8151	<.0001
		Covariance		98.0729	<.0001	206.16	<.0001	245.22	<.0001
Model Fit		Deviance		238258		231176		227071	
Statistics		AIC		238276		231194		227089	
		BIC		238338		231256		227152	

Model E9 - Math		Parameter		Grade 3		Grade 8		Grade 10	
Fixed Effect	Initial Status	Intercept	γ_{00}	594.29	<.0001	774.02	<.0001	807.53	<.0001
		Gender	γ_{01}	3.4628	0.0107	-.3538	0.7515	3.4795	0.0012
		SES	γ_{02}	53.7703	<.0001	37.0995	<.0001	31.7637	<.0001
	Rate of change	Time	γ_{10}	39.1672	<.0001	15.9586	<.0001	9.9895	<.0001
		SES*Time	γ_{11}	-.3955	0.4454	4.3139	<.0001	3.8672	<.0001
Variance	Level 1	Within Person		874.31	<.0001	577.52	<.0001	619.33	<.0001
Component	Level 2	In Initial Status		3797.44	<.0001	2343.38	<.0001	1993.21	<.0001
		In Rate of Change		179.51	<.0001	67.969	<.0001	43.0469	<.0001
		Covariance		109.86	<.0001	228.54	<.0001	274.22	<.0001
Model Fit		Deviance		301950		274962		266136	
Statistics		AIC		301968		274980		266154	
		BIC		302033		275044		266218	

Model E10 - Math		Parameter		Grade 3		Grade 8		Grade 10	
Fixed Effect	Initial Status	Intercept	γ_{00}	594.32	<.0001	773.62	<.0001	807.32	<.0001
		Gender	γ_{01}	3.5963	0.0095	-.2364	0.8327	3.4805	0.0012
		SES	γ_{02}	53.5604	<.0001	37.8268	<.0001	35.1522	<.0001
	Rate of change	Time	γ_{10}	39.1239	<.0001	18.3934	<.0001	12.1121	<.0001
		Gender*Time	γ_{11}	-.2326	0.6487	-.6954	0.0824	-.00958	0.9811
Variance	Level 1	Within Person		874.31	<.0001	577.54	<.0001	619.35	<.0001
Component	Level 2	In Initial Status		3797.44	<.0001	2343.43	<.0001	1993.18	<.0001
		In Rate of Change		179.53	<.0001	72.4857	<.0001	46.7354	<.0001
		Covariance		109.85	<.0001	227.79	<.0001	273.87	<.0001
Model Fit		Deviance		301951		275076		266227	
Statistics		AIC		301969		275094		266245	
		BIC		302033		275158		266309	

Table 14a. Estimated and Empirical Mean Scores

Model A		Model B				Model C1	Female			Male		
Grade	Total	Grade	M1	M2	M3	Grade	M1	M2	M3	M1	M2	M3
3	657	3	618	657	696	3	616	655	695	620	659	698
8	810	8	792	810	828	8	792	810	828	791	809	828
10	839	10	827	839	851	10	824	837	849	829	841	853
Empirical		3	619	657	695	3	617	655	693	621	658	696
		8	792	809	828	8	792	809	829	792	809	827
		10	827	839	851	10	824	837	849	829	841	853
Model A Sample1		Model B Sample-1				Model C2	Low-SES			Not Low-SES		
Grade	Total	Grade	M1	M2	M3	Grade	M1	M2	M3	M1	M2	M3
3	670	3	630	670	709	3	596	635	674	649	689	728
8	820	8	801	820	839	8	774	790	806	815	831	847
10	848	10	835	848	861	10	809	819	829	845	855	865
Empirical						3	597	634	673	650	689	725
						8	774	790	806	812	830	852
						10	810	818	830	840	856	868
Model A Smaple-2		Model B Sample-2				Model C3	Hispanic/Latino			White		
Grade	Total	Grade	M1	M2	M3	Grade	M1	M2	M3	M1	M2	M3
3	659	3	620	659	697	3	594	634	675	640	681	721
8	809	8	791	809	827	8	779	794	810	810	825	841
10	838	10	826	838	850	10	814	825	836	841	852	863
Empirical						3	594	633	673	641	680	718
						8	778	795	809	806	824	846
						10	814	825	836	837	853	864
						Model C4	African American			White		
						Grade	M1	M2	M3	M1	M2	M3
						3	589	626	663	643	681	718
						8	769	785	800	810	825	840
						10	805	814	824	842	852	861
Empirical						3	590	625	661	641	680	718
						8	769	785	800	806	824	846
						10	806	813	824	837	853	864

Table 14b. Estimated and Empirical Mean Scores

Model	Grade	Female Hispanic/Latino			Female White			Male Hispanic/Latino			Male White		
		M1	M2	M3	M1	M2	M3	M1	M2	M3	M1	M2	M3
	3	591	631	672	637	678	718	596	637	677	643	683	724
D1	8	778	794	810	809	825	841	778	794	810	809	825	841
	10	812	822	833	838	849	838	817	828	839	844	854	865
	3	591	631	670	637	678	715	597	635	675	644	683	721
Empirical	8	779	795	811	806	823	846	778	794	808	807	824	846
	10	815	824	835	834	850	861	813	827	837	840	856	868
Model	Grade	Female. African American			Female White			Male African American			Male White		
	3	587	624	662	642	679	716	590	627	665	645	682	719
D2	8	770	785	801	810	826	842	768	784	800	809	824	840
	10	803	812	822	840	850	859	807	816	825	844	853	863
	3	591	625	662	637	678	715	589	624	661	644	683	721
Empirical	8	771	787	803	806	823	846	767	782	797	807	824	846
	10	805	814	825	834	850	861	806	812	823	840	856	868
Model	Grade	Low-SES H/L			Low-SES White			Not Low-SES H/L			Not Low-SES White		
	3	589	629	670	618	658	699	626	666	707	654	695	735
D3	8	773	788	803	788	802	817	807	822	837	822	837	852
	10	808	819	829	823	833	844	834	845	855	849	859	870
	3	589	628	669	619	656	697	627	664	701	655	696	732
Empirical	8	774	790	806	784	801	819	796	814	826	818	836	860
	10	811	822	832	821	832	843	824	836	849	844	862	873
Model	Grade	AA Low-SES			AA Not Low-SES			White Low-SES			White Not Low-SES		
	3	580	618	655	617	655	692	621	658	695	658	695	732
D4	8	761	776	790	789	807	821	790	805	819	822	836	850
	10	797	805	814	822	830	839	826	834	842	851	859	867
	3	582	616	653	618	654	692	619	656	697	655	696	732
Empirical	8	764	778	793	783	801	818	784	801	819	818	836	860
	10	800	805	816	816	827	839	821	832	843	844	862	873
Model	Grade	Female Low-SES			Female Not Low-SES			Male Low-SES			Male Not Low-SES		
	3	594	634	673	648	687	726	598	637	676	651	690	730
D5	8	774	790	807	815	832	848	773	789	806	814	831	847
	10	808	818	828	843	853	863	811	821	831	847	857	867
	3	596	633	672	646	686	723	598	635	673	653	692	728
Empirical	8	775	791	807	811	829	852	773	788	804	812	831	852
	10	810	818	829	837	853	865	810	819	830	843	859	871

Table 14c. Estimated and Empirical Mean Scores

Model	Grade	Female Hispanic/Latino			Female White			Male Hispanic/Latino			Male White		
		M1	M2	M3	M1	M2	M3	M1	M2	M3	M1	M2	M3
E1	3	591	631	671	638	678	718	597	637	677	644	683	723
	8	778	797	816	805	824	844	778	797	817	805	825	844
	10	812	824	837	836	848	861	817	830	843	841	854	867
Empirical	3	591	631	670	637	678	715	597	635	675	644	683	721
	8	779	795	811	806	823	846	778	794	808	807	824	846
	10	815	824	835	834	850	861	813	827	837	840	856	868
Model	Grade	Hispanic/Latino Female			Hispanic/Latino Male			White Female			White Male		
E2	3	591	631	671	597	637	677	637	678	718	643	683	724
	8	778	794	809	779	794	810	809	825	840	810	825	841
	10	812	823	834	817	828	839	838	849	860	843	854	865
Empirical	3	591	631	670	597	635	675	637	678	715	644	683	721
	8	779	795	811	778	794	808	806	823	846	807	824	846
	10	815	824	835	813	827	837	834	850	861	840	856	868
Model	Grade	Female African American			Female White			Male African American			Male White		
E3	3	586	625	664	640	679	717	589	628	667	643	682	720
	8	769	788	806	806	824	843	768	786	805	805	823	841
	10	803	815	827	836	848	860	806	818	830	840	852	864
Empirical	3	591	625	662	637	678	715	589	624	661	644	683	721
	8	771	787	803	806	823	846	767	782	797	807	824	846
	10	805	814	825	834	850	861	806	812	823	840	856	868
Model	Grade	African American Female			African American Male			White Female			White Male		
E4	3	587	624	662	590	627	665	642	679	716	645	682	719
	8	770	785	800	769	784	799	810	826	841	809	825	840
	10	803	812	822	807	816	825	840	850	859	844	853	863
Empirical	3	591	625	662	589	624	661	637	678	715	644	683	721
	8	771	787	803	767	782	797	806	823	846	807	824	846
	10	805	814	825	806	812	823	834	850	861	840	856	868
Model	Grade	H/L Low-SES			H/L Not Low-SES			White Low-SES			White Not Low-SES		
E5	3	589	629	670	626	667	707	618	658	698	655	695	736
	8	773	788	804	805	820	836	789	804	820	821	836	852
	10	808	819	830	832	843	854	824	835	846	847	858	870
Empirical	3	589	628	669	627	664	701	619	656	697	655	696	732
	8	774	790	806	796	814	826	784	801	819	818	836	860
	10	811	822	832	824	836	849	821	832	843	844	862	873

Model	Grade	Low-SES H/L			Low-SES White			Not Low-SES H/L			Not Low-SES White		
	3	589	629	670	618	658	699	626	666	707	654	695	735
E6	8	772	789	806	785	806	819	808	825	841	820	837	854
	10	808	820	831	822	836	844	835	846	857	848	859	870
	3	589	628	669	619	656	697	627	664	701	655	696	732
Empirical	8	774	790	806	784	801	819	796	814	826	818	836	860
	10	811	822	832	821	832	843	824	836	849	844	862	873
Model	Grade	AA Low-SES			AA Not Low-SES			White Low-SES			White Not Low-SES		
	3	580	618	655	618	655	692	621	658	695	658	695	732
E7	8	761	776	792	790	805	820	791	806	822	820	835	850
	10	797	806	816	819	829	838	826	836	845	849	858	868
	3	582	616	653	618	654	692	619	656	697	655	696	732
Empirical	8	764	778	793	783	801	818	784	801	819	818	836	860
	10	800	805	816	816	827	839	821	832	843	844	862	873
Model	Grade	Low-SES AA			Low-SES White			Not Low-SES AA			Not Low-SES White		
	3	580	618	657	619	657	696	617	656	694	656	695	733
E8	8	761	777	792	788	803	819	793	809	825	820	835	851
	10	797	806	816	823	833	842	823	832	842	849	858	868
	3	582	616	653	619	656	697	618	654	692	655	696	732
Empirical	8	764	778	793	784	801	819	783	801	818	818	836	860
	10	800	805	816	821	832	843	816	827	839	844	862	873
Model	Grade	Low-SES Female			Low-SES Male			Not Low-SES Female			Not Low-SES Male		
	3	594	633	673	598	637	676	648	687	726	651	690	729
E9	8	774	790	806	774	790	806	815	831	847	815	831	847
	10	808	818	828	811	821	831	843	853	863	847	857	867
	3	596	633	672	598	635	673	646	686	723	653	692	728
Empirical	8	775	791	807	773	788	804	811	829	852	812	831	852
	10	810	818	829	810	819	830	837	853	865	843	859	871
Model	Grade	Female Low-SES			Female Not Low-SES			Male Low-SES			Male Not Low-SES		
	3	594	633	673	648	687	726	598	637	676	651	690	729
E10	8	774	792	810	811	830	848	773	791	809	811	829	847
	10	807	819	832	839	852	864	811	823	835	843	855	867
	3	596	633	672	646	686	723	598	635	673	653	692	728
Empirical	8	775	791	807	811	829	852	773	788	804	812	831	852
	10	810	818	829	837	853	865	810	819	830	843	859	871

Table 15. Summary of Effects of Level-1 and Level-2 Predictors

Model	Tested Effects	GR 3	GR 8	GR 10
Model A	Means	Y	Y	Y
A1		Y	Y	Y
A2		Y	Y	Y
Model B	Time (fixed)	Y	Y	Y
B1		Y	Y	Y
B2		Y	Y	Y
Model C1	Gender	N	N	N
C2	SES	N	Y	Y
C3	Race1	N	Y	Y
C4	Race2	Y	Y	Y
Model D1	Gender/Race1	N/N	N/Y	N/Y
D2	Gender/Race2	N/Y	N/Y	N/Y
D3	SES/Race1	N/N	Y/Y	Y/N
D4	SES/Race2	N/Y	Y/Y	Y/Y
D5	Gender/SES	N/Y	N/Y	N/Y
Model E1	Gender * Race1	N	N	N
E2	Race1 * Gender	N	Y	Y
E3	Gender * Race2	N	0.08	N
E4	Race2 * Gender	Y	Y	Y
E5	Race1 * SES	N	Y	Y
E6	SES * Race1	0.05	Y	Y
E7	Race2 * SES	Y	Y	Y
E8	SES * Race2	N	Y	Y
E9	SES * Gender	N	Y	Y
E10	Gender * SES	N	0.08	N

Y indicates statistically significant effects ($p<.0001$)
N indicates that the statistical significant level is less than .10.