

**Academy School  
District Twenty**

**Effects of IB Participation on Reading Achievement and Growth, 2000-2004**

**in**

**Academy School District Twenty**

**Submitted to**

**Dr. Ken Vedra, Superintendent  
Dr. Ken Turner, Deputy Superintendent**

**Dr. Vonda L. Kiplinger  
Director for Research and Planning  
Academy School District Twenty**

**June 9, 2005**

## Executive Summary

This report describes the results of the longitudinal study of the effects of the International Baccalaureate (IB) program on reading achievement and growth from 2001 through 2004 in grades 5-8, and 8-10. Although the IB program is an “open-enrollment” program, admitting any student who requests admittance, IB students consistently out-perform their peers on all CSAP assessments. The question that this study seeks to answer is whether the higher academic achievement is most likely due to selection effects such as academic ability and prior achievement, motivation, family background characteristics, etc., or whether the IB program has a *unique*, value-added effect on student achievement and growth in reading.

Several different analytic techniques are used to answer the research questions, including descriptive procedures, t-tests for differences in means, bivariate correlations, multiple linear regression, and the multilevel mixed effects technique, hierarchical linear modeling (HLM). Hierarchical Linear Models (HLM) is used to evaluate the effects of IB participation on student growth in mathematics achievement, controlling for student background characteristics. The HLM models estimated are models of individual change.

The results of the models are mixed.

- ✚ IB students consistently out-perform non-IB students in reading.
- ✚ Analyses of middle school cohort (i.e., grades 4-8) indicate that this higher performance may be attributable to selection effects and/or to the effects of other variables not available for analysis, but analyses of the youngest and oldest cohorts (grades 3-5 and 7-10) indicate that for students scoring at the same level in grade 7, IB may have a slightly positive effect on reading growth rate during the middle and early high school years.
- ✚ One argument for the most likely explanation of IB students’ high performance is that they are brighter and more motivated. However, the results also indicate the possibility of an accrual of benefits from participation in the IB program over time. The analysis of the effect of length of time in the IB program during 7<sup>th</sup> through 10<sup>th</sup> grades indicates that for each year of participation in IB, reading growth rate increases by about one and a half scale score points per year, for an approximate additional boost of six and one-half points for the student who participates all four years. To put this apparent increase in perspective, however, the reader must remember that, although statistically significant, the total accrual effect appears to be less than one percent.
- ✚ The lack of definitive evidence of positive effects of IB participation on reading achievement and growth should *not* be viewed as negative evidence. The homogeneity and typically high performance of all three cohorts of students make it difficult to discriminate between groups of students and evaluate the effectiveness of many school programs on achievement. Add to that (a) the inability to measure many variables hypothesized to be related to achievement, (b) self-selection and program effectiveness (including achievement prior to the measurements included in this study, IB participation during earlier grades, and family SES), and (c) the inefficiency of the proxy variables available for analysis, estimation of true program effects is not as clear nor as precise as we would wish.

## Table of Contents

Table of Contents.....	ii
1. Introduction .....	1
2. Research Questions .....	1
3. Methodology .....	2
3.1 Descriptive.....	2
3.2 Multiple Regression and Correlation .....	2
3.3 Hierarchical Linear Models (HLM).....	3
3.3.1 Outcomes Models.....	4
3.3.2 Growth Models.....	5
4. Results .....	6
4.1 Results for the 2002-2004 Grade 3-5 Cohort.....	7
4.1.1 Description of the 2002-2004 Grade 3-5 Elementary School Cohort .....	7
4.1.2 Results from the Multivariate Analyses.....	12
4.1.3 Results from the HLM Estimation of the Effects of IB Participation on Reading.....	16
Achievement in Elementary School .....	16
4.1.4 Summary of Results for the 2002-2004 Grade 3-5 Cohort .....	25
4.2 Results for the 2000-2004 Grade 4-8 Cohort.....	26
4.2.1 Description of the 2000-2004 Grade 4-8 Cohort.....	26
4.2.2 Results from the Multivariate Analyses.....	32
4.2.3 Results from the HLM Estimation of the Effects of IB Participation on Reading.....	36
Achievement and Growth in Middle School .....	36
4.2.4 Summary of Results for the 2000-2004 Grade 4-8 Cohort .....	45
4.3 Results for the 2001-2004 Grade 7-10 Cohort.....	45
4.3.1 Description of the 2001-2004 Grade 7-10 Secondary School Cohort .....	45
4.3.2 Results from the Multivariate Analyses.....	50
4.3.3 Results from the HLM Estimation of Achievement Growth.....	55
5. Conclusions .....	65

## List of Tables

- Table 1. Analytical Cohorts by Year and Grade
- Table 2. Dependent Variables by Cohort
- Table 3. Number of Years of Participation in District 20's Primary Years IB Program
- Table 4. Mean Reading Achievement and Other Characteristics of IB and Non-IB Students, Grade 3-5, CSAP 2002-2004
- Table 5. Effects of Initial Reading Status and IB Participation on Reading Achievement CSAP, Grade 5, 2004
- Table 6. Results from the IB Main Effects Model
- Table 7. Effects of IB Participation on Reading Growth in Grades 3-5, 2002-2004
- Table 8. Number of Years of Participation in District 20's IB Program, 2000-2004 Grade 4-8 Student Cohort
- Table 9. Mean Reading Scores of IB and Non-IB Students, Grades 4-8, 2000-2004
- Table 10. Comparison of Unconditional Means and Unconditional Growth Models
- Table 11. Comparison of IB Main Effects Model and a Value Added Model
- Table 12. Number of Years of Participation in District 20's IB Program, 2000-2004 Grade 4-8 Student Cohort
- Table 13. Mean Reading Scores of IB and Non-IB Students, Grades 7-8, 2001-2004
- Table 14. Comparison of Unconditional Means and Unconditional Growth Models 2001-2004
- Table 15. Effects of IB Participation on Reading Growth in Grades 7-10, 2001-2004
- Table 16. Effects of Length of IB Participation on Reading Growth during Grades 7-10, 2001-2004

## List of Figures

- Figure 1a. Distribution of Reading Scores on the Grade 3 CSAP, 2002 for Non-IB Students
- Figure 1b. Distribution of Reading Scores on the Grade 3 CSAP, 2002 for IB Students
- Figure 2a. Distribution of Reading Scores on the Grade 5 CSAP, 2004 for Non-IB Students
- Figure 2b. Distribution of Reading Scores on the Grade 5 CSAP, 2004 for IB Students
- Figure 3a. Nonparametric Reading Growth Curves, Grades 3-5, 2002-2004 for a Sample of IB Students
- Figure 3b. Nonparametric Reading Growth Curves, Grades 3-5, 2002-2004 for a Sample of Non-IB Students
- Figure 4a. Ordinary Least Squares Reading Growth Trajectories, Grades, 3-5, 2002-2004 for a Sample of IB Students
- Figure 4b. Ordinary Least Squares Reading Growth Trajectories, Grades, 3-5, 2002-2004 for a Sample of Non-IB Students
- Figure 5a. Mean OLS Regression Line for Reading Growth, Grades 3-5, 2002-2004 for a Sample of IB Students
- Figure 5b. Mean OLS Regression Line for Reading Growth, Grades 3-5, 2002-2004 for a Sample of Non-IB Students
- Figure 6. Reading Growth Trajectories of IB and Non-IB Students, Grades 3-5, 2002-2004
- Figure 7. Effect of IB Participation on Reading Growth Rate, Grades 3-5, 2002-2004
- Figure 8a. Distribution of Reading Scores on the Grade 4 CSAP, 2000 for Non-IB Students
- Figure 8b. Distribution of Reading Scores on the Grade 4 CSAP, 2000 for IB Students
- Figure 9a. Distribution of Reading Scores on the Grade 8 CSAP, 2004 for Non-IB Students
- Figure 9b. Distribution of Reading Scores on the Grade 8 CSAP, 2004 for Students
- Figure 10a. Nonparametric Reading Growth Curves, Grades 4-8, 2000-2004 for a Sample of IB Students
- Figure 10b. Nonparametric Reading Growth Curves, Grades 4-8, 2000-2004 for a Sample of Non-IB Students
- Figure 11a. Ordinary Least Squares Reading Growth Trajectories, Grades 4-8, 2000-2004 for a Sample of IB Students
- Figure 11b. Ordinary Least Squares Reading Growth Trajectories, Grades 4-8, 2000-2004 for a Sample of Non-IB Students
- Figure 12a. Mean OLS Regression Line for Reading Growth, Grades 4-8, 2000-2004 for a Sample of IB Students
- Figure 12b. Mean OLS Regression Line for Reading Growth, Grades 4-8, 2000-2004 for a Sample of Non-IB Students
- Figure 13. Reading Growth Trajectories of IB and Non-IB Students, Grades 4-8, 2000-2004
- Figure 14. Effect of IB Participation on Reading Growth Rate, Grades 4-8, 2000-2004
- Figure 15a. Distribution of Reading Scores on the Grade 7 CSAP, 2001 for Non-IB Students
- Figure 15b. Distribution of Reading Scores on the Grade 7 CSAP, 2001 for IB Students
- Figure 16a. Distribution of Reading Scores on the Grade 10 CSAP, 2004 for Non-IB Students
- Figure 16b. Distribution of Reading Scores on the Grade 10 CSAP, 2004 for IB Students

- Figure 17a. Nonparametric Reading Growth Curves, Grades 7-10, 2001-2004 for a Sample of IB Students
- Figure 17b. Nonparametric Reading Growth Curves, Grades 7-10, 2001-2004 for a Sample of Non-IB Students
- Figure 18a. Ordinary Least Squares Reading Growth Trajectories, Grades 7-10, 2001-2004 for a Sample of IB Students
- Figure 18b. Ordinary Least Squares Reading Growth Trajectories, Grades 7-10, 2001-2004 for a Sample of Non-IB Students
- Figure 19a. Mean OLS Regression Line for Reading Growth, Grades 7-10, 2001-2004 for a Sample of IB Students
- Figure 19b. Mean OLS Regression Line for Reading Growth, Grades 7-10, 2001-2004 for a Sample of Non-IB Students
- Figure 20. Effect of IB Participation on Reading Growth Rate, Grades 7-10, 2001-2004
- Figure 21. Reading Growth Trajectories of IB and Non-IB Students, Grades 7-10, 2001-2004

# Effects of IB Participation on Reading Achievement and Growth, 2000-2004

## 1. Introduction

This report is the first in a series of reports that describe the results of a multi-component study to examine the effects of IB, AP, and Honors program participation on student achievement and growth in student achievement in reading and mathematics at the elementary and secondary levels. This report describes the results of the longitudinal study of the effects of the International Baccalaureate (IB) program on reading achievement and growth from 2000 through 2004 in grades 3-5, 4-8, and 7-10<sup>1</sup>. The elementary and secondary study cohorts are overlapping in grades, but not individuals, since students in grades 3 through 10 during the period 2000-2004 are examined. Data are not available prior to 2000 and, in some cases 2001, depending on the cohort.

This report tells the “reading story” in Academy School District Twenty from 2000 through 2004. Described are the primary research questions, methodology, and results of the study of the effects of participation in the IB program in elementary and secondary schools on student achievement and growth in reading achievement (as measured by CSAP) from 2000-2004. IB students consistently out-perform their peers on all CSAP assessments. The question that this study seeks to answer is whether the higher academic achievement is most likely due to selection effects such as academic ability and prior achievement, motivation, family background characteristics, etc., or whether the IB program has a *unique*, value-added effect on student achievement and growth.

## 2. Research Questions

This phase of the research project investigates three primary research questions related to the participation of secondary school students in the IB program. The research questions are:

1. Adjusting for student background characteristics, what is the effect of IB on student reading achievement (outcomes) and growth?
2. Do the effects of IB vary by school level (elementary, middle, and high)?
3. Is there an interaction between student characteristics and the effects of the IB program on reading achievement? In other words, is the IB program more effective for some students than for others?
4. Adjusting for student background characteristics, what is the effect of length of time in the IB program on student reading achievement (outcomes) and growth?

---

<sup>1</sup> Forthcoming reports will detail the effects of the IB program on math achievement and growth, 2001-2004, grades 5-8 and 2002-2004, grades 8-10; effects of AP courses, 2002-2004 grades 8-10; and effects of the Honors program, 2002-2004, grades 8-10. School District 20

### 3. Methodology

Several different analytic techniques are used to answer the research questions and to provide context for inferences from those results. The analytic techniques include descriptive procedures, t-tests for differences in means, bivariate correlations, multiple linear regression, and hierarchical linear modeling (HLM).

#### 3.1 Descriptive

Prior to conducting multivariate analyses and hierarchical modeling, descriptive statistics were examined. Frequency distributions, crosstabulation tables, and t-tests (for differences in means) compare IB students and non-IB students. The variables compared include CSAP reading, length of time in the IB program, gender, minority, poverty status (defined by free or reduced lunch eligibility), language background, and other demographics. In addition, the raw growth trajectories of a sample of students are examined in to order to determine the best parametric fit (e.g., linear or polynomial) to the actual data and to visually inspect the variation in empirical growth curves for IB and non-IB students.

#### 3.2 Multiple Regression and Correlation

In order to identify factors that are potentially predictive of student achievement and growth in reading, bivariate correlations and stepwise multiple linear regressions are completed prior to estimating the hierarchical linear models.

The dependent (criterion) variables are the Spring 2004 CSAP scores in reading and average gain scores from the 2000, 2001, or 2002 through 2004 CSAP reading assessments, depending on the cohort. As discussed above, the cohorts are overlapping in grades and time, but not in years. Table 1 lists the cohorts and their grades by year<sup>2</sup>:

Table 1  
Analytical Cohorts by Year and Grade

Cohort	Year/Grade				
	2000	2001	2002	2003	2004
2002-2004 Grade 3-5 Cohort <sup>3</sup>			3	4	5
2000-2004 Grade 4-8 Cohort	4	5	6	7	8
2001-2004 Grade 7-10 Cohort		7	8	9	10

<sup>2</sup> Year refers to the Spring CSAP administrations.

<sup>3</sup> Cohort years refer to the times of the Spring CSAP administrations. Thus, 2002 refers to the 2001-2002 school year.

Table 2 lists the dependent variables in the multiple regression and correlation analyses for each cohort:

Table 2  
Dependent Variables by Cohort

Cohort	Dependent Variables	
2002-2004 Grade 3-5 Cohort	2004 Grade 5 CSAP Reading Score	Avg. Gain in Reading 2002-2004
2000-2004 Grade 4-8 Cohort	2004 Grade 8 CSAP Reading Score	Avg. Gain in Reading 2000-2004
2001-2004 Grade 7-10 Cohort	2004 Grade 10 CSAP Reading Score	Avg. Gain in Reading 2001-2004

Potential independent variables (predictors) include IB participation, length of time in IB, prior achievement in reading (as measured by the initial CSAP assessment listed above), gender, minority status, poverty status, language background, cognitive disability, and non-cognitive disability. The possible interaction of gender and IB participation also was examined. Interaction effects are of interest to determine if program (IB) participation has a differential effect on the achievement of males and females. It is impossible to test interaction of SES and minority status with IB participation because there are no low SES students in the secondary IB program and very few minority students. However, this is not surprising, given that this is a high SES-low minority district.

Before proceeding to more complex analyses, empirical, nonparametric growth plots of IB and non-IB students in each cohort are examined. It is always desirable to examine the “raw” empirical growth plots in order to visualize how students change over time and to determine the best overall parametric model fit (e.g., linear, logistic) to the data. Random samples of 25 IB and 25 non-IB students in each cohort are selected and their nonparametric growth plots examined. Empirical growth plots allow us to evaluate change in reading performance in both absolute terms (i.e., against the overall scale) and in relative terms (relative to other students’ growth) as well as to visually inspect the variation in growth curves for IB and non-IB students.

In addition, ordinary least squares (OLS) regression models are fit to the data in order to determine whether there is enough variability in the data to warrant further analyses.

### 3.3 Hierarchical Linear Models (HLM)

The multilevel mixed effects technique, Hierarchical Linear Models (HLM), is used to evaluate the effects of IB participation and length of participation on student achievement and on growth in achievement, controlling for student background characteristics. Two types of HLM models are estimated: outcomes models and individual change models. While HLM is the method of choice for both types of mixed models, the procedures and model specifications are fundamentally different (Kiplinger, 2004). Outcomes models are similar to standard multiple regression models except that separate regression equations are estimated for each level of the hierarchy. Estimation of the effects on individual change over time in the multilevel context is known collo-

quially as growth modeling. The basic model specifications for the reading multilevel analyses are described below.

### 3.3.1 Outcomes Models

The outcomes models are 2-level achievement models where Level 1 specifies the student effects and level 2 evaluates program effects on achievement outcomes. The dependent variables are student scores on the 2004 CSAP reading assessments.

Potential Level 1 predictors include prior achievement in reading, gender, minority, poverty status, language background, and cognitive disability. The Level 2 predictor is IB status.

An example of a 2-level outcomes model is illustrated below. Level 1 is the student level model and Level 2 is the programmatic level model.

#### **LEVEL 1 MODEL**

$$READ04_{ij} = \beta_{0j} + \beta_{1j}(READ02_{ij} - \overline{READ02_{..}}) + r_{ij}$$

#### **LEVEL 2 MODEL**

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(IB\_SCHOO_j) + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}(IB\_SCHOO_j) + u_{1j}$$

Substituting the Level 2 coefficients into the Level 1 equation, we arrive at the combined model:

$$READ04_{ij} = \gamma_{00} + \gamma_{01} * IB\_SCHOO_j + \gamma_{10} * (READ02_{ij} - \overline{READ02_{..}}) + \gamma_{11} * IB\_SCHOO_j * (READ02_{ij} - \overline{READ02_{..}}) + u_{0j} + u_{1j} * (READ02_{ij} - \overline{READ02_{..}}) + r_{ij}$$

where

READ04 is the reading score for student *i* in school *j* in 2004.

READ02 is initial status, reading score in 2002. Initial status is centered around the grand mean for all schools, which can be interpreted as an adjusted mean for school *j*. In this case the intercept,  $\beta_{0j}$ , is the expected outcome for a student in school *j* whose score is at the mean of all schools (grand mean).

$\beta_{0j}$  = Adjusted mean outcome in school *j* after controlling for differences in initial reading status

$\beta_{1j}$  = Main effect of initial reading status on reading achievement

$\gamma_{00}$  = Adjusted mean reading achievement for non-IB students in 2004

$\gamma_{01}$  = Main effect of IB participation

$\gamma_{10}$  = Main effect of initial status

$\gamma_{11}$  = Cross-level interaction of initial status in reading and IB participation

$r_{ij}$  = Unique person effect

$u_{0j}$  = Unique school effect on the intercept (mean school outcome after controlling for differences in initial reading status and IB participation)

$u_{1j}$  = Unique school effect on the slope (effect of initial reading status on reading achievement after controlling for differences IB participation).

### 3.3.2 Growth Models

The growth models are 2-level models, where Level 1 specifies the individual growth trajectory (based on CSAP scores), which is modeled as dependent on person level factors, including IB status. These person-level effects are modeled in Level 2 as the slope coefficients for achievement growth. Initially, 3-level models also were attempted in which school effects were included in level 3, but these models could not be estimated. Thus, all growth models in this study are 2 level models.

The dependent variable in the growth models is student score on the 2004 CSAP reading assessments. In the 2-level growth models Level 1 specifies the individual growth trajectory (based on CSAP scores), and contains student test scores and indicators of test score patterns. The individual growth trajectories are then modeled as dependent on person level factors. These person-level effects are modeled in Level 2 as the intercept and slope coefficients for achievement growth. Potential Level 2 predictors include prior achievement in reading, gender, minority, poverty status, language background, cognitive disability, IB status, and time in the IB program.

Prior to estimation of IB effects on reading achievement and growth, preliminary unconditional models are estimated. The first unconditional model is an unconditional *means* model, i.e., a model with no predictors in either level 1 or level 2. This model estimates the total amount of true variation in the outcome and allows us to initially partition the total variation in the outcome that is between and within persons. Initially partitioning the variance components provides an indication of whether there is enough variability within persons to warrant growth modeling and whether there is enough variability between persons to warrant a predictive model at level 2. The second unconditional model is an unconditional *growth* model, a model with time as the only level 1 predictor and no level 2 predictors. This type of model indicates whether we can account for the within-person variance by modeling growth alone or whether additional variability between persons in intercepts and slopes is great enough to model with level 2 predictors.

An example of a 2-level growth model is illustrated below. The level-1 specifies the growth trajectory and level 2 models the effects of individual factors, including IB status. This model examines effects of IB participation on reading growth rate without controlling for other student characteristics. The Level 1 and 2 submodels are:

## LEVEL 1 MODEL

$$\text{READING}_{it} = \pi_{0i} + \pi_{1i} (\text{YEAR}_{it}) + e_{it}$$

## LEVEL 2 MODEL

$$\pi_{0i} = \beta_{00} + r_{0i}$$

$$\pi_{1i} = \beta_{10} + \beta_{11} (\text{IB}) + r_{1i}$$

Substituting the Level 2 coefficients into the Level 1 equation, we arrive at the combined model:

$$\text{READING}_{it} = \beta_{00} + \beta_{10} (\text{YEAR}) + \beta_{11} (\text{IB*YEAR}) + r_{0i} + r_{1i} (\text{YEAR}_{it}) + e_{it}$$

where,

$\text{READING}_{it}$  is the outcome variable, reading score in 2004

$\pi_{0i}$  = Initial reading status of person  $i$ , that is, the expected outcome for that student in the spring of Year = 0

$\pi_{1i}$  = Rate of change in reading (growth rate) for person  $i$

YEAR is 0 at the initial measurement, 1 at the second measurement, 2 at the third measurement, and so on

$\beta_{00}$  = Estimated mean intercept, or initial reading status of students at year 0

$\beta_{10}$  = Mean academic year growth rate in reading

$\beta_{11}$  = Mean effect of IB on reading growth rate

IB = 0 if student was in IB program for less than 2 years or never in the program or 1 if student was in the IB program for 2 or more years

$e_{it}$  = Level 1 residual variance in true growth trajectory of person  $i$  (within-persons deviation)

$r_{0i}$  = Level 2 residual variance in true intercept across all individual in the population

$r_{1i}$  = Level 2 residual variance in true rate of change across all individual in the population (between-persons deviation)

## 4. Results

This chapter describes the results of the descriptive, multivariate, and HLM analyses of the three cohorts:

- 2002-2004 Grade 3-5 Cohort
- 2000-2004 Grade 4-8 Cohort
- 2001-2004 Grade 7-10 Cohort

The temporal boundaries of these cohorts are determined by the CSAP schedule of assessments and the availability of data on IB participation in the district.

## 4.1 Results for the 2002-2004 Grade 3-5 Cohort

Results of the descriptive, multivariate, and hierarchical linear models (HLM) analyses are discussed below.

### 4.1.1 Description of the 2002-2004 Grade 3-5 Elementary School Cohort

There were 962 students in the 2002-2004 Grade 3-5 Cohort, 14 percent of whom (131) participated at least two years in the primary years IB program during 2002-2004. Table 3 indicates that IB participation ranged from 1 to 3 years during the primary grades in 2002 – 2004, with the vast majority participating 3 years.

Table 3  
Number of Years of Participation in District 20's Primary Years IB Program  
2002 – 2004 Grade 3-5 Cohort

Years in IB Program	n	Percent
1	7	5.3
2	16	12.2
3	108	82.5
Total	131	100

Although the results of the descriptive analyses indicate that IB students score significantly higher than non-IB students in reading (659 vs. 648 in 2004), there appears to be no difference in their growth rates. IB students do not differ from other students with respect to initial status on the 2002 Grade 3 reading assessment, minority status, language background, poverty status, or cognitive disabilities. These results are shown in Table 4.

Table 4  
Mean Reading Achievement and Other Characteristics  
of IB and Non-IB Students, Grade 3-5  
CSAP 2002 – 2004

Characteristic	IB Students	Non-IB Students
Mean 5 <sup>th</sup> Grade Reading Score, 2004	659	648
Mean 3 <sup>rd</sup> Grade Reading Score, 2002	608	606
Percent Female	55	48
Percent Poverty	2	5
Percent Minority	8	13

Characteristic	IB Students	Non-IB Students
Percent non-English Language	4	1
Percent with a Cognitive Disability	3	3
No. of Students	131	831

The distributional characteristics of IB/non-IB reading achievement indicate that both non-IB and IB students span the range of reading scores. Possible reading scores on the vertical scale for the grade 3 through 10 CSAP range from 300 to 900. Figures 1a-b and 2a-b illustrate the distributions on the 2002 grade 3 and 2004 grade 5 CSAP reading assessments, respectively.

These figures indicate that the grade 3 reading scores of non-IB students range from 408 to 795, while the range for IB students is 406 to 795, and that there is no difference in their means.

Figure 1a  
Distribution of Reading Scores on the  
Grade 3 CSAP, 2002 for  
Non-IB Students

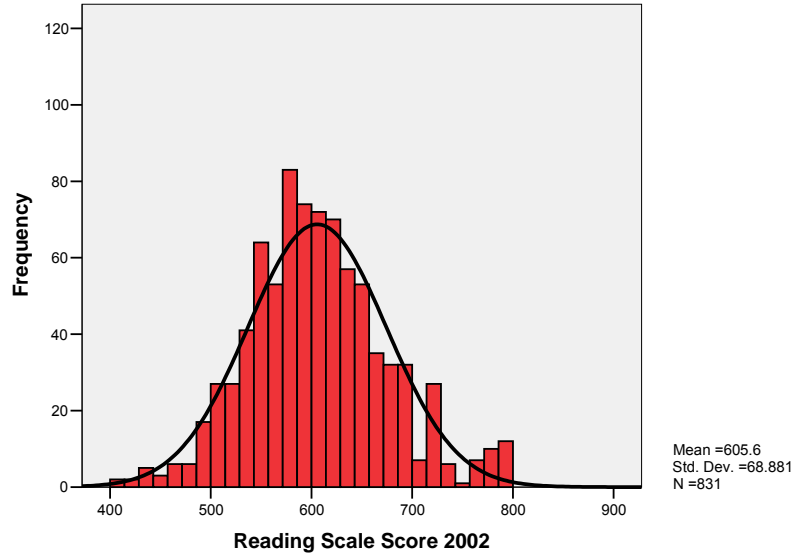
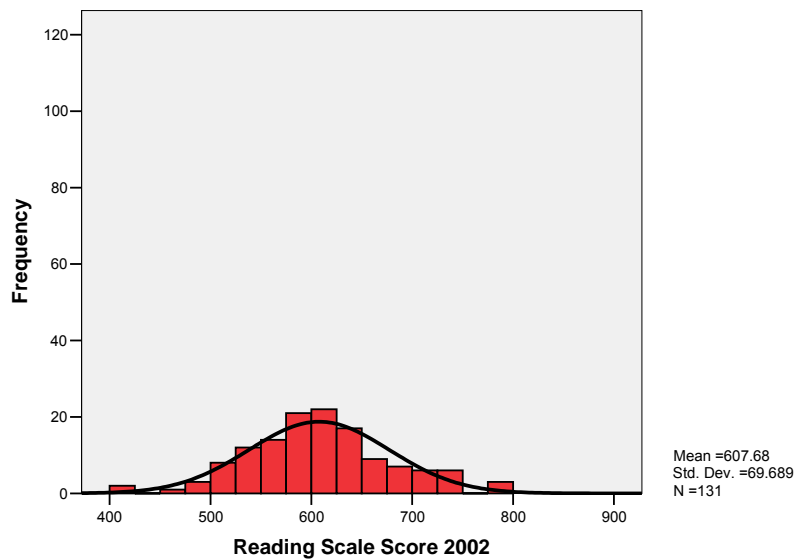


Figure 1b  
Distribution of Reading Scores on the  
Grade 3 CSAP, 2002 for  
IB Students



The distributions of scores on the 2004 grade 5 reading assessment are similar to those discussed above, with both IB and non-IB students spanning the range of scores and no obvious concentration on scores (e.g., in the upper tail for IB students). The distributions of the grade 5 reading scores are illustrated in Figures 2a-2b below.

Figure 2a  
 Distribution of Reading Scores on the  
 Grade 5 CSAP, 2004 for  
 Non-IB Students

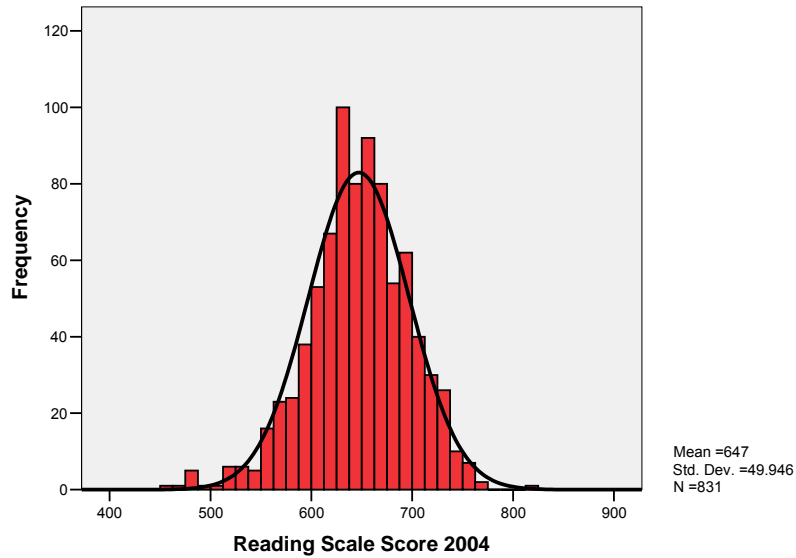
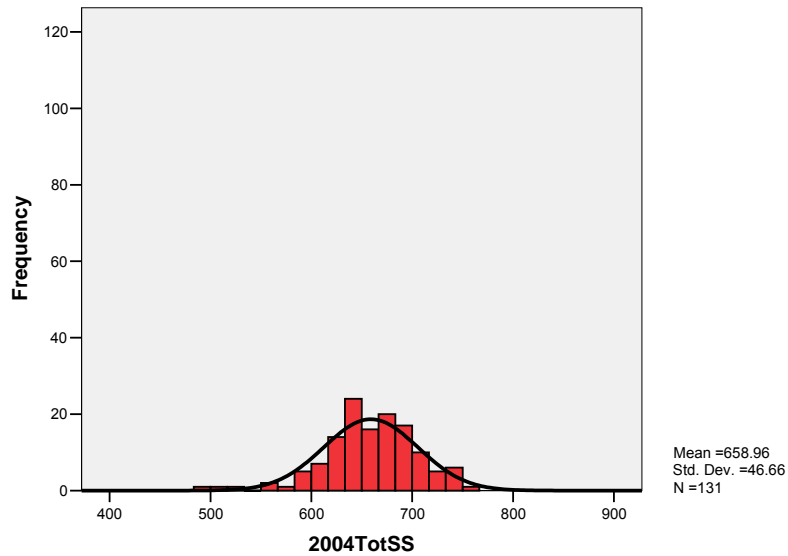


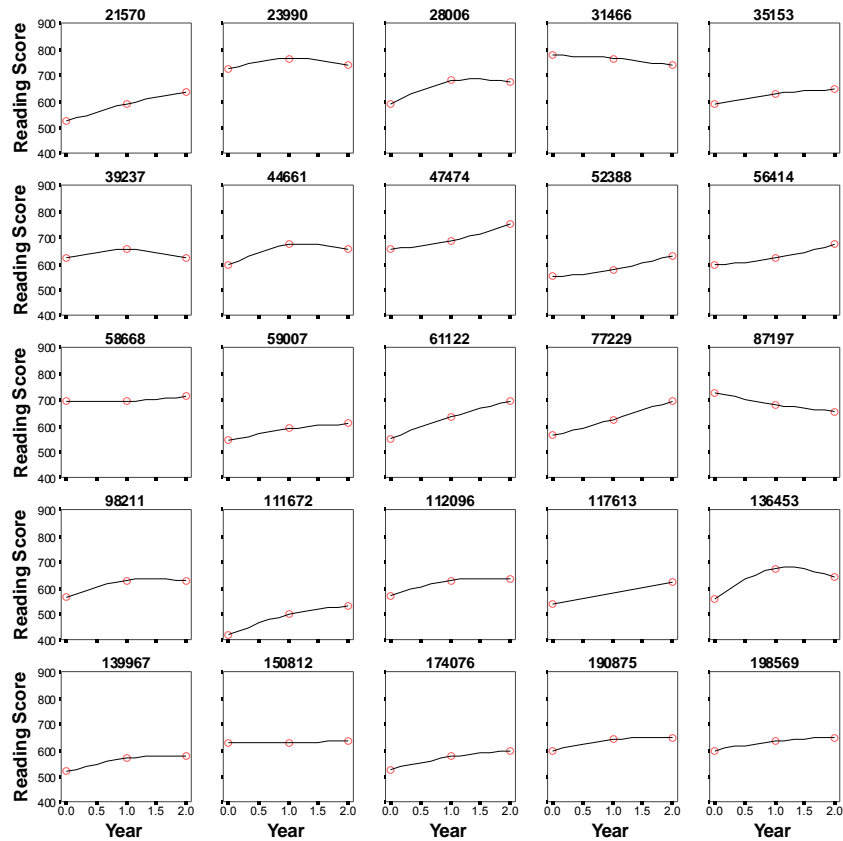
Figure 2b  
 Distribution of Reading Scores on the  
 Grade 5 CSAP, 2004 for  
 IB Students



The “raw” empirical growth plots are examined in order to visualize how this cohort of students change over time and to determine the best overall parametric model fit (e.g., linear, logistic) to the data. Random samples of 25 IB and 25 non-IB students are selected and their empirical, nonparametric, growth plots examined.

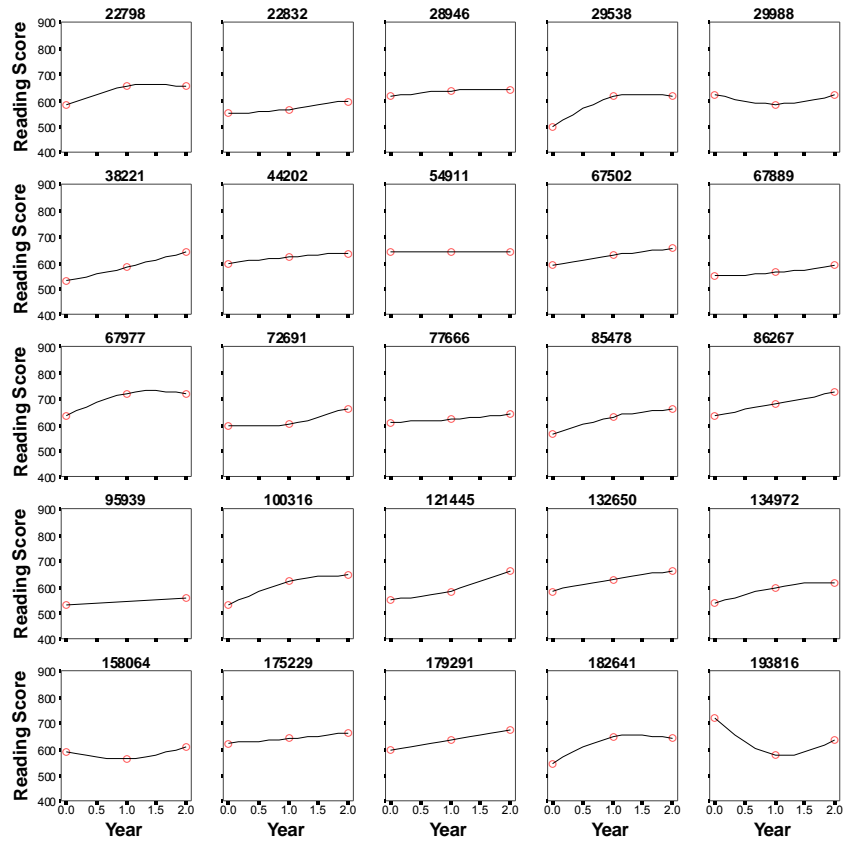
The “smoothed” empirical growth curves are illustrated in Figures 3a and 3b for IB and non-IB students, respectively.

Figure 3a  
 Nonparametric Reading Growth Curves  
 Grades 3-5, 2002-2004 for a  
 Sample of IB Students<sup>4</sup>



<sup>4</sup> The “id numbers” above each student plot are assigned at random and cannot be used to identify individual students.

Figure 3b  
 Nonparametric Reading Growth Curves  
 Grades 3-5, 2002-2004 for a  
 Sample of Non-IB Students<sup>5</sup>



In these two samples of 25 students, it is apparent that the basic shapes of the curves for both groups are similar although there are notable exceptions (e.g., 44661 and 136453 in the IB sample and 29538 and 193816 in the non-IB sample). The best common functional form across individual trajectories appears to be linear.

#### 4.1.2 Results from the Multivariate Analyses

Results from the ordinary least squares fits to the “raw” data and from the exploratory multiple regression predictive models are discussed in this section.

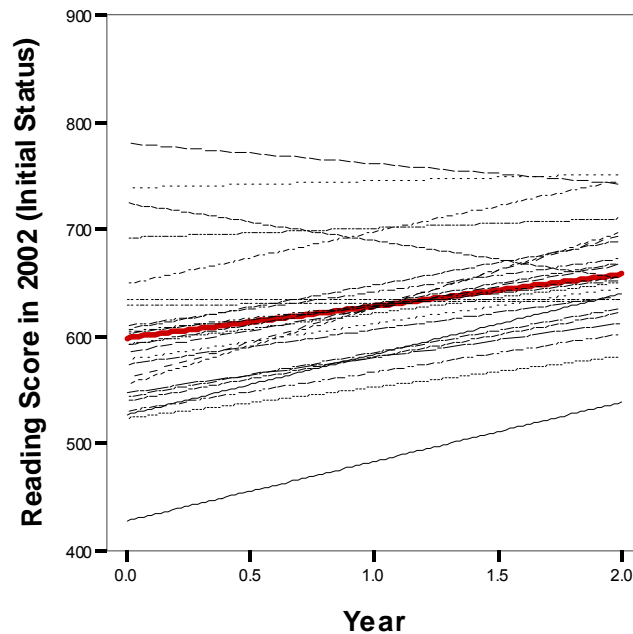
<sup>5</sup> The “id numbers” above each student plot are assigned at random and cannot be used to identify individual students.

#### 4.1.2.1 Results from the Ordinary Least Squares Fits

The ordinary least squares (OLS) fit to the raw achievement data indicates sufficient variability to warrant more sophisticated modeling. The OLS growth trajectories of the samples of 25 IB and 25 non-IB students are shown in Figures 4a and 4b, respectively.

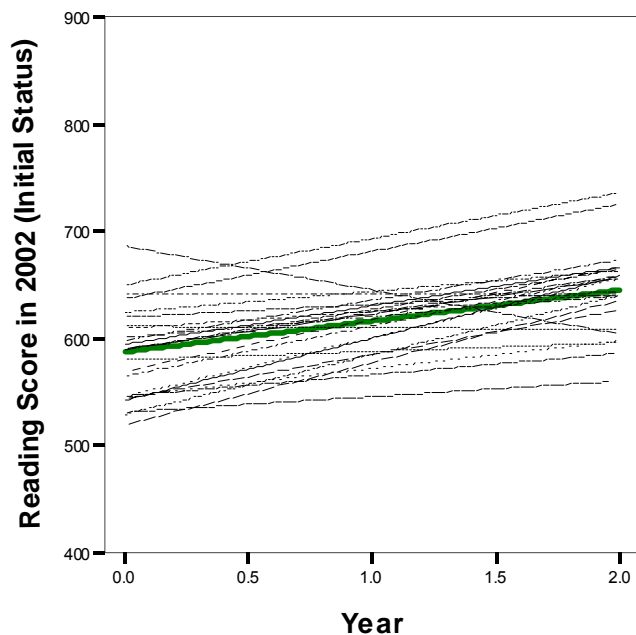
Figures 4a and 4b indicate that there is substantially more variability in the intercepts (2002 grade 3 test scores) of IB students than for non-IB students. The figures also indicate that, with a few exceptions, the slopes (growth rates) of both groups of students are similar. The red line in Figure 4a and the green line in Figure 4b represent the average regression lines for the two groups of students. Although visual inspection of the two figures seems to indicate slightly higher mean initial status (reading score in Year 0) and growth rate in the IB sample the differences are not statistically significant.

Figure 4a  
Ordinary Least Squares Reading Growth Trajectories  
Grades 3-5, 2002-2004 for a Sample of IB Students



Legend: Black lines indicate individual growth trajectories.  
The red line is the average regression line for IB students.

Figure 4b  
Ordinary Least Squares Reading Growth Trajectories  
Grades 3-5, 2002-2004 for a Sample of Non-IB Students



Legend: Black lines indicate individual growth trajectories.  
The green line is the average regression line for non-IB students.

Figures 5a and 5b illustrate the mean regression line (center line) and the 95% confidence interval around the regression line (indicated by the two outer lines) for the samples of IB and non-IB students. The red circles represent individual reading scores.

Figure 5a  
 Mean OLS Regression Line for Reading Growth  
 Grades 3-5, 2002-2004 for a Sample of IB Students

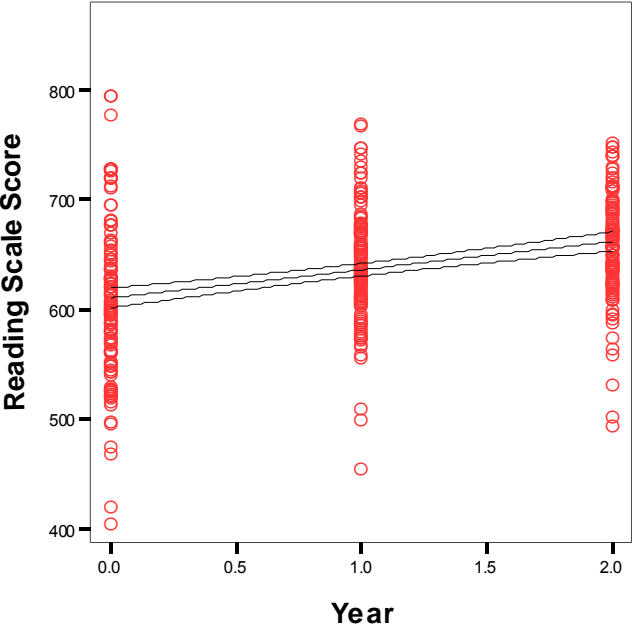
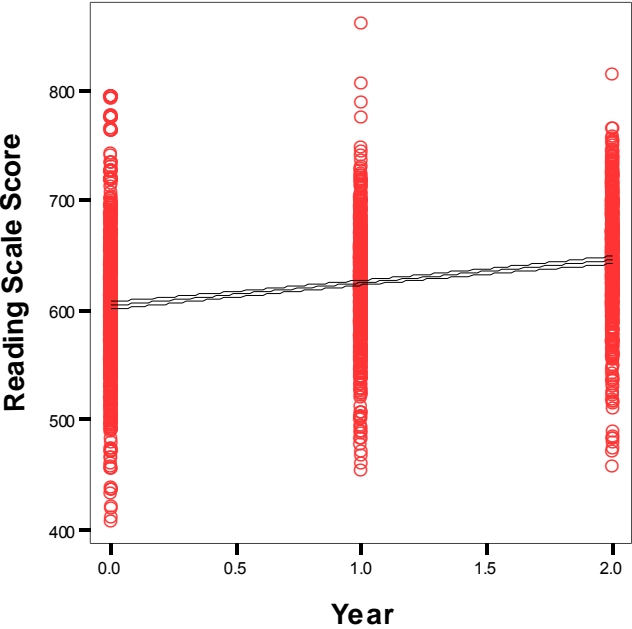


Figure 5b  
 Mean OLS Regression Line for Reading Growth  
 Grades 3-5, 2002-2004 for a Sample of Non-IB Students



The OLS regressions indicate no difference in mean grade 3 reading assessments (Year 0) of IB and non-IB students (608 and 606, respectively). The rate of change (slope) is also similar for IB and non-IB students (21 vs. 26 points per year, respectively).

#### 4.1.2.2 Results from the Multiple Regression: Exploratory Predictive Models

In order to identify factors that are potentially predictive of student achievement and growth in reading, bivariate correlations and stepwise linear regression models are estimated prior to the hierarchical linear modeling. The bivariate correlations indicate that the strongest correlate to the outcome, reading achievement in third grade, and average gain in achievement from third through fifth grades is reading score in grade 3 ( $r = 0.69$  and  $-0.69$ , respectively).

The preliminary multiple regression analyses indicate no significant effect of IB participation on either reading achievement or growth in elementary school. The next section describes the results of the hierarchical linear models (HLM) of reading achievement and growth during elementary school in District 20.

#### 4.1.3 Results from the HLM Estimation of the Effects of IB Participation on Reading Achievement in Elementary School

The OLS fits to the raw achievement data discussed in § 4.1.2.1 indicate sufficient variability to warrant more sophisticated modeling. Hierarchical Linear Models (HLM) is used to examine the effects of IB participation on reading growth and outcomes. The results of the HLM models are described below.

##### 4.1.3.1 Results from the HLM Outcomes Model

The preliminary multiple regression analyses indicate no significant effect of IB participation on either reading achievement or growth in elementary school. Results of the HLM analyses are consistent with these conclusions. When the IB predictor variable is introduced into the student level model, either alone or in combination with initial status, the “effect” is not significant. The only significant predictor of achievement outcomes is initial reading status (i.e., reading performance on the Grade 3 CSAP reading assessment in 2002). These results indicate that at the elementary level, the IB program does not make a *unique* contribution to student reading achievement.

The specification for this model is described in § 3.3.1 of the Methodology section. To aid in understanding and interpreting the results, the model equations and specifications are repeated below:

**LEVEL 1 MODEL**

$$READ04_{ij} = \beta_{0j} + \beta_{1j}(READ02_{ij} - \overline{READ02_{..}}) + r_{ij}$$

**LEVEL 2 MODEL**

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(IB\_SCHOO_j) + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}(IB\_SCHOO_j) + u_{1j}$$

Substituting the Level 2 coefficients into the Level 1 equation, we arrive at the combined model:

$$READ04_{ij} = \gamma_{00} + \gamma_{01} * IB\_SCHOO_j + \gamma_{10} * (READ02_{ij} - \overline{READ02_{..}}) + \gamma_{11} * IB\_SCHOO_j * (READ02_{ij} - \overline{READ02_{..}}) + u_{0j} + u_{1j} * (READ02_{ij} - \overline{READ02_{..}}) + r_{ij}$$

where

READ04 indicates the grade 5 reading score for student *i* in school *j* in 2004.

READ02 is initial status, or the grade 3 reading score in 2002. Initial status is centered around the grand mean for all schools, which can be interpreted as an adjusted mean for school *j*. In this case the intercept,  $\beta_{0j}$ , is the expected outcome for a student in school *j* whose score is at the mean of all schools (grand mean).

$\beta_{0j}$  = Adjusted mean outcome in school *j* after controlling for differences in initial reading status

$\beta_{1j}$  = Main effect of initial reading status on reading achievement in 2004

$\gamma_{00}$  = Adjusted mean reading achievement for non-IB students

$\gamma_{01}$  = Main effect of IB participation

$\gamma_{10}$  = Main effect of initial status

$\gamma_{11}$  = Cross-level interaction of initial status in reading and IB participation

$r_{ij}$  = Unique person effect

$u_{0j}$  = Unique school effect on the intercept (mean school outcome after controlling for differences in initial reading status and IB participation)

$u_{1j}$  = Unique school effect on the slope (effect of initial reading status on reading achievement after controlling for differences IB participation).

Table 5 displays the results for the fixed and random effects.

Table 5  
Effects of Initial Reading Status and IB Participation on Reading Achievement  
CSAP, Grade 5, 2004

		Parameter	Coefficient
<b>Fixed Effects</b>			
Initial status,		$\gamma_{00}$	647.80***
	$\beta_{0i}$		(2.01)
		$\gamma_{01}$	10.59**
Rate of change,		$\gamma_{10}$	0.50***
	$\beta_{1i}$		(0.02)
		$\gamma_{11}$	- 0.05
			(0.06)
<b>Variance Components</b>			
Level 1	Within- persons	$r_{ti}$	1224.10
Level 2	For initial status	$\mu_{0i}$	33.53***
	For rate of change	$\mu_{1i}$	0.00
	Covariance	$\mu_{01}$	0.06
	Correlation	$r_{01}$	0.30
Goodness-of-fit Deviance			9586.52

Numbers in parentheses represent standard errors.

\*\*\*  $p < .001$

\*\*  $p > .05$  and  $\leq .10$

Thus, the mean score for non-IB students on the 2004 Grade 5 CSAP reading assessment was 648 ( $\gamma_{00}$ ). The main effect of initial status ( $\gamma_{10}$ ) is 0.5, which means that, on average, a one-point increment in initial status is associated in a one-half point increment in scores on the 2004 reading assessment. The results for IB participation ( $\gamma_{01}$ ) indicate that at the elementary level the program does not contribute significantly to reading achievement. Likewise, there is no interaction between initial status and IB participation ( $\gamma_{11}$ ).

Decomposition of the explained variance indicates that there is substantial variation left unexplained by this model. The variance components (random effects) are displayed in Table 5.

The intercept variance of 33.53 indicates significant ( $p = 0.001$ ) variability among schools in terms of their average reading achievement. Regarding the slopes, the  $X^2$  significance test indicates that no significant variation in the initial status-2004 achievement regression slopes re-

mains unexplained after controlling for initial status. The largest variance component (1224.1) is at level-1 of the model, (students within schools), indicating that substantial within school variation remains unexplained by this model.

Comparing the residual variance of this model with the total variance indicated by the unconditional model indicates that amount of variation that is explained by the final model. The variance obtained by the unconditional model is the overall variability in the true school means. The percentage reduction in that variance is the percentage of true variance that is explained by the model. Comparison of these models indicate that inclusion of initial status as a predictor of grade 5 reading achievement “explains” approximately two-thirds of the variability among schools in terms of average reading achievement (33.53 vs. 98.08 for the unconditional model). Inclusion of initial status “explains” nearly one-half of the level-1 (within persons) variability in reading achievement (1224.10 vs. 2374.19 for the unconditional model).

Between schools variability:  $((98.08 - 33.53) / 98.0) = 0.66$

Within persons variability:  $((2374.19 - 1224.10) / 2374.19) = 0.48$ .

When the total variability in reading achievement is considered, this model explains only about half of the total variability in fifth grade reading achievement.

% of total variance explained by model =

$$\frac{\text{total variability from unconditional model} - \text{residual variance from final model}}{\text{total variability from unconditional model}}$$

$$= \frac{2472.27 - 1257.63}{2472.27} = .49$$

#### 4.1.3.2 Effects of IB Participation on Reading Growth in Elementary School

The preliminary regression analyses indicate that none of the available variables, including initial reading status and IB participation, are significant predictors of reading growth in these elementary schools. The HLM analyses confirm these results. A series of 2-level growth models are estimated using HLM. In multilevel models of individual change (i.e., growth modeling), Level 1 specifies the individual growth trajectory, which is modeled as dependent on person-level factors. These person-level effects are modeled in Level 2 as the slope coefficients for achievement growth. The dependent, or criterion, variable is student score on the grade 8 2004 CSAP reading assessment.

Although the unconditional growth model indicates sufficient variation in both intercept and slope to warrant more complex modeling, subsequent models indicate that neither IB participation nor initial status “explain” the observed variation in reading growth. Goodness-of-fit statistics indicate that models including IB participation are no better than the unconditional growth model. The results of the analysis of the main effects of IB participation are provided in Table 6.

This model examines effects of IB participation on the intercept and slope without controlling for other student characteristics. The Level 1 and 2 submodels are:

**LEVEL 1 MODEL**

$$\text{READING}_{it} = \pi_{0i} + \pi_{1i} (\text{YEAR}_{it}) + e_{it}$$

**LEVEL 2 MODEL**

$$\pi_{0i} = \beta_{00} + \beta_{01} (\text{IB}) + r_{0i}$$

$$\pi_{1i} = \beta_{10} + \beta_{11} (\text{IB}) + r_{1i}$$

Substituting the Level 2 coefficients into the Level 1 equation, we arrive at the combined model:

$$\text{READING}_{it} = \beta_{00} + \beta_{01} (\text{IB}) + \beta_{10} (\text{YEAR}) + \beta_{11} (\text{IB} * \text{YEAR}) + r_{0i} + r_{1i} (\text{YEAR}_{it}) + e_{it}$$

where,

READING<sub>it</sub> is the outcome variable, reading score in grade 5

$\pi_{0i}$  = Initial reading status of person *i*, that is, the expected outcome for that student in the spring of grade 3 (when Year = 0)

$\pi_{1i}$  = Rate of change in reading (growth rate) for person *i*

YEAR is 0 at spring 2002, 1 at spring 2003 and 2 at Spring 2004 (grades 3, 4, and 5)

$\beta_{00}$  = Estimated mean intercept, or initial reading status of a non-IB student at year 0

$\beta_{01}$  = Mean IB/non-IB gap on the intercept (initial reading status)

$\beta_{10}$  = Mean academic year growth rate in reading

$\beta_{11}$  = Mean effect of IB on reading growth rate

IB = 0 if student was in IB program for less than 2 years or never in the program or 1 if student was in the IB program for 2 or more years

$e_{it}$  = Level 1 residual variance in true growth trajectory of person *i* (within-persons deviation)

$r_{0i}$  = Level 2 residual variance in true intercept across all individual in the population

$r_{1i}$  = Level 2 residual variance in true rate of change across all individual in the population (between-persons deviation)

The fixed and random effects from this model are provided in Table 6.

Table 6  
Results from the IB Main Effects Model

		Parameter	Coefficient
<b>Fixed Effects</b>			
Initial status,		$\beta_{00}$	604.12*** (2.211)
$\pi_{0i}$		$\beta_{01}$	6.36 <i>ns</i> (6.11)
Rate of change,		$\beta_{10}$	20.70*** (0.861)
$\pi_{1i}$		$\beta_{11}$	4.95 <i>ns</i> (2.275)
<b>Variance Components</b>			
Level 1	Within- persons	$e_{ij}$	700.85
Level 2	For initial status	$r_{0i}$	3513.430***
	For rate of change	$r_{1i}$	261.803***
Goodness-of-fit			
Deviance			29690.65
$\chi^2$ statistic			23.98 <i>ns</i>

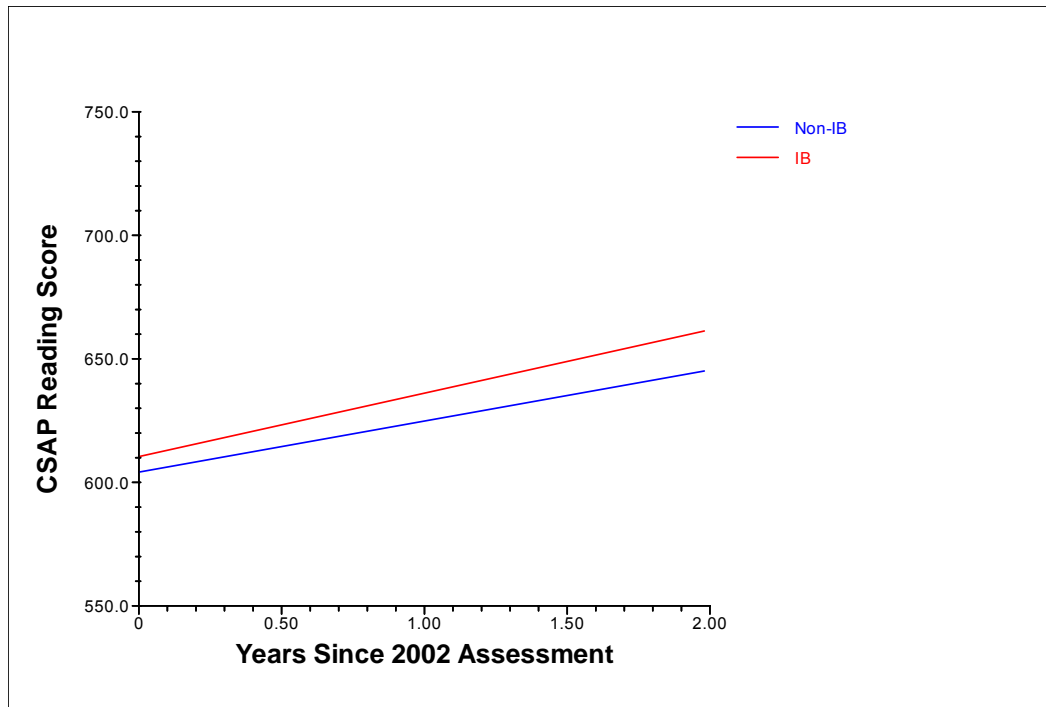
Numbers in parentheses represent standard errors.

\*\*\*  $p < .001$  *ns* not statistically significant

On average, students in this cohort scored about 605 points ( $\beta_{00}$ ) on the 2002 grade 3 CSAP reading assessment and gained an average of 21 points ( $\beta_{10}$ ) annually from grade 3 through grade 5. There are no significant differences in either initial status (score on the grade 3 assessment) or reading growth rates of IB and non-IB students ( $\beta_{01}$  and ( $\beta_{11}$ , respectively).

The growth trajectories of IB and non-IB students are compared in Figure 6 below.

Figure 6  
 Reading Growth Trajectories of IB and Non-IB Students  
 Grades 3-5, 2002-2004



Although it appears that IB students score slightly higher than non-IB students in elementary school and that their growth rate is also slightly higher, the differences are not statistically significant.

Since IB and non-IB students demonstrate the same average reading score on the 2002 grade 3 assessment, a different model in which IB status was excluded from the model of the intercept (i.e., initial status) was estimated<sup>6</sup>. This model, described below, examines the effects of IB participation on reading growth rate without controlling for initial status or other student characteristics. If IB participation during elementary school has a positive effect on reading achievement and growth, then we would expect to see a steeper slope for IB students among those who start at the same initial status, in this case the district average on the grade 4 reading assessment. The Level 1 and 2 submodels are:

<sup>6</sup> Also, we obviously cannot attribute higher initial performance to subsequent IB participation. In effect, this is how the equations are specified in the HLM level 2 model presented on page 20.

## LEVEL 1 MODEL

$$\text{READING}_{it} = \pi_{0i} + \pi_{1i} (\text{YEAR}_{it}) + e_{it}$$

## LEVEL 2 MODEL

$$\pi_{0i} = \beta_{00} + r_{0i}$$

$$\pi_{1i} = \beta_{10} + \beta_{11} (\text{IB}) + r_{1i}$$

Substituting the Level 2 coefficients into the Level 1 equation, we arrive at the combined model:

$$\text{READING}_{it} = \beta_{00} + \beta_{10} (\text{YEAR}) + \beta_{11} (\text{IB} * \text{YEAR}) + r_{0i} + r_{1i} (\text{YEAR}_{it}) + e_{it}$$

where,

READING<sub>it</sub> is the outcome variable, reading score in grade 5

$\pi_{0i}$  = Initial reading status of person *i*, that is, the expected outcome for that student in the spring of grade 3 (when Year = 0)

$\pi_{1i}$  = Rate of change in reading (growth rate) for person *i*

YEAR is 0 at spring 2002, 1 at spring 2003 and 2 at spring 2004 (grades 3, 4, and 5)

$\beta_{00}$  = Estimated mean intercept, or initial reading status of students at year 0

$\beta_{10}$  = Mean academic year growth rate in reading

$\beta_{11}$  = Mean effect of IB on reading growth rate

IB = 0 if student was in IB program for less than 2 years or never in the program or 1 if student was in the IB program for 2 or more years (during 3<sup>rd</sup> - 5<sup>th</sup> grades)

$e_{it}$  = Level 1 residual variance in true growth trajectory of person *i* (within-persons deviation)

$r_{0i}$  = Level 2 residual variance in true intercept across all individual in the population

$r_{1i}$  = Level 2 residual variance in true rate of change across all individual in the population (between-persons deviation)

In this model, the intercept ( $\beta_{00}$ ) is the grand mean for initial status (grade 3 reading mean score) across all students. The coefficient on the slope (rate of change) represents the effect of IB participation on reading growth rate ( $\beta_{10}$ ). Students in this cohort scored, on average, about 605 points on the 2002 grade 3 CSAP reading assessment ( $\beta_{00}$ ) and averaged a growth rate of about 20 scale score points per year ( $\beta_{10}$ ) from third through seventh grades. The results suggest that IB students' reading growth rate is slightly higher than the average growth rate by nearly 7 points ( $\beta_{11}$ ,  $p < .001$ ). That is, IB students' average reading growth rate is about 27 points, as compared to the overall average of 20 points.

The variance decomposition is provided in Table 7. Unfortunately, it appears that the addition of IB participation explains very little of the residual between-person variance in rate of change ( $r_{1i}$  = 255 for the unconditional growth model, and  $r_{1i}$  = 263 in the conditional model, a reduction of only 3%).

Table 7  
Effects of IB Participation on Reading Growth in Grades 3 – 5  
2002 – 2004

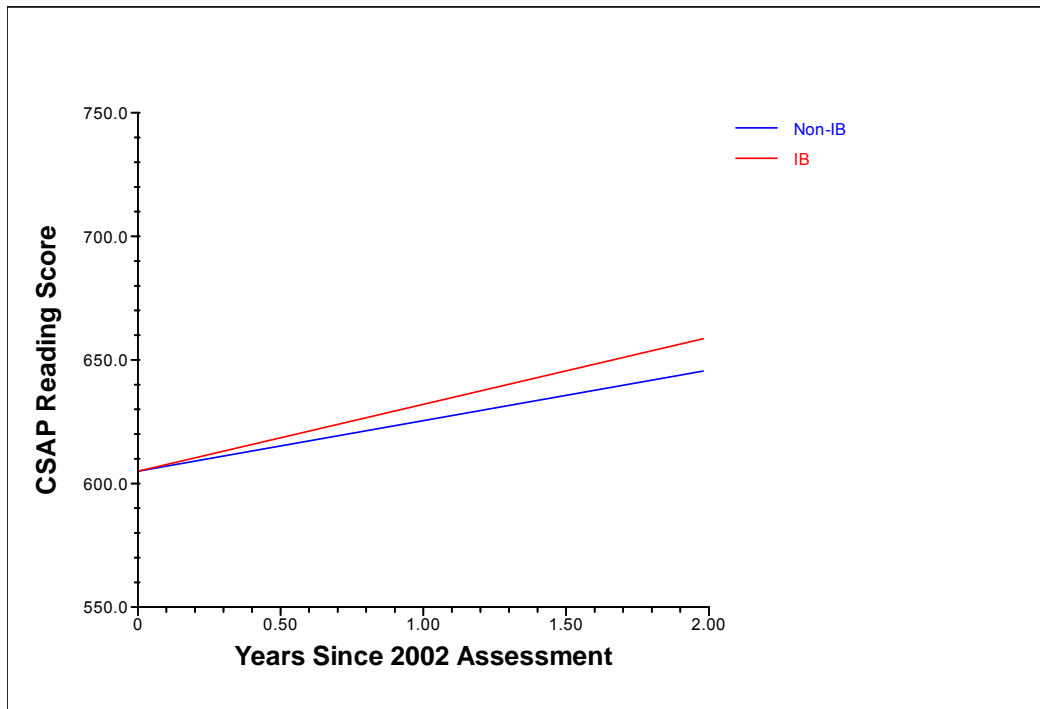
		Parameter	Coefficient
<b>Fixed Effects</b>			
Initial status,		$\beta_{00}$	604.99*** (2.063)
$\pi_{0i}$			
Rate of change,		$\beta_{10}$	20.47*** (0.836)
$\pi_{1i}$		$\beta_{11}$	6.59*** (1.604)
<b>Variance Components</b>			
Level 1	Within- persons	$e_{it}$	700.12***
Level 2	For initial status	$r_{0i}$	3515.23***
	For rate of change	$r_{1i}$	262.878***
	Covariance	$r_{01}$	-713.98
	Correlation	$r_{01}$	- 0.743
Goodness-of-fit Deviance			29699.03

Numbers in parentheses represent standard errors.

\*\*\*  $p < .001$

These results are illustrated in Figure 7.

Figure 7  
Effect of IB Participation on Reading Growth Rate  
Grades 3-5, 2002-2004



This plot depicts the average growth trajectories of IB and non-IB students. The intercept is the district average of 605 scale score points on the grade 3 reading assessment in 2002. The growth trajectories illustrate the small, but significant, effect of IB participation during elementary school on reading achievement and growth. Starting from the district average, IB students appear to have about a six and one-half point advantage in growth rate. The more common approach in to illustrate the growth trajectories of IB and non-IB students, projected from their respective average intercepts, as seen in Figure 6. Figure 6 depicts the estimated growth trajectories of students who score at the average of their groups, IB and non-IB students. The value in considering the information illustrated by Figure 7 is that controlling for starting point, IB students appear to progress from grade 3 through 5 at a slightly faster rate than non-IB students.

#### 4.1.4 Summary of Results for the 2002-2004 Grade 3-5 Cohort

In summary, the analyses reported above indicate no difference in the average grade 3 reading scores and growth rates of IB and non-IB students during elementary school. However, the last analysis reported above suggest that for students scoring at the same level in grade 3, IB may have a slightly positive effect on reading growth rate during elementary school.

## 4.2 Results for the 2000-2004 Grade 4-8 Cohort

Results of the descriptive, OLS, multivariate, and HLM analyses for the 2000-2004 grade 4-8 cohort are discussed below.

### 4.2.1 Description of the 2000-2004 Grade 4-8 Cohort

There were 893 students in this cohort, 21 percent (191 students) of whom have been or are currently enrolled in the IB program in District 20 during elementary and/or middle school. As shown in Table 8 below, IB participation ranged from 1 year to 5 years during the period of the 1999-2000 through 2003-2004 school years. Most students in the program, 70 percent, had participated two or three years. Only three percent had participated only 1 year, while 27 percent had participated four or five years.

Table 8  
Number of Years of Participation in District 20's IB Program  
2000-2004 Grade 4-8 Student Cohort

Years in IB Program	n	Percent
1	6	3.0
2	68	35.6
3	66	34.7
4	4	2.1
5	47	24.6
Total	191	100

In District 20, IB and non-IB students do not differ with respect to demographics such as gender, minority status, and cognitive or non-cognitive disabilities. However, IB students in this cohort are significantly more likely to have a non-English language background than non-IB students (3 % vs. 0 %, respectively) and to come from more advantaged homes<sup>7</sup> (1 % vs. 4%, respectively, from poverty homes).

Students in the IB program consistently out-perform students not in the program, as shown in Table 9 below. However, there is no significant difference between these subgroups of students in their rate of growth in reading from fourth through eighth grades.

---

<sup>7</sup> Student eligibility for free or reduced price lunch serves as the proxy for poverty status. "Advantaged homes" are those whose children are not eligible for free or reduced price lunch.

Table 9  
Mean Reading Scores<sup>a</sup> of IB and Non-IB Students, Grades 4-8, 2000-2004

Reading Assessment	IB Students	Non-IB Students	Mean Difference	<i>P</i> value
2000 Grade 4	644	612	32	<.001
2001 Grade 5	663	636	27	<.001
2002 Grade 6	667	644	23	<.001
2003 Grade 7	686	662	24	<.001
2004 Grade 8	703	677	26	<.001
Average Annual Gain, 2000-2004	15	16	1	.101

<sup>a</sup> Average scaled scores on the CSAP reading assessments.

However, the distributional characteristics of the IB/non-IB achievement indicate that both groups of students span the range of scores on the CSAP, as illustrated in Figures 8 and 9.

Figure 8a  
Distribution of Reading Scores on the  
Grade 4 CSAP, 2000 for  
Non-IB Students

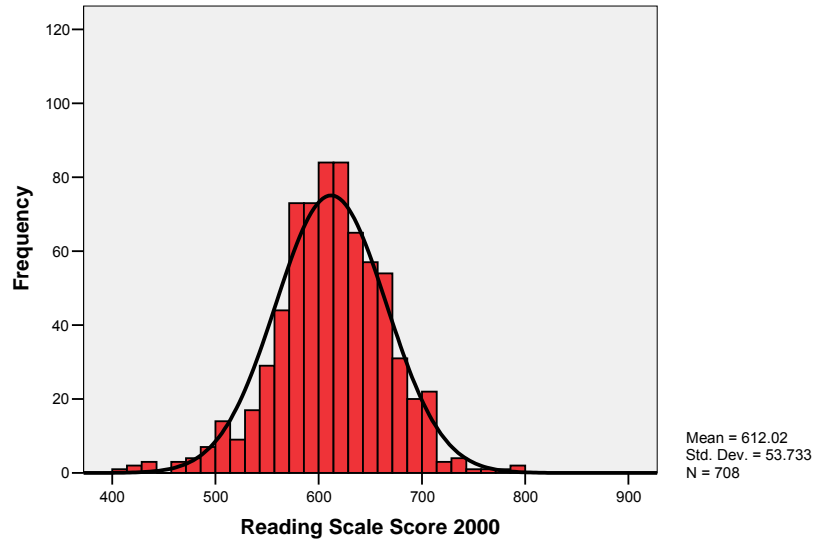


Figure 8b  
Distribution of Reading Scores on the  
Grade 4 CSAP, 2000 for  
IB Students

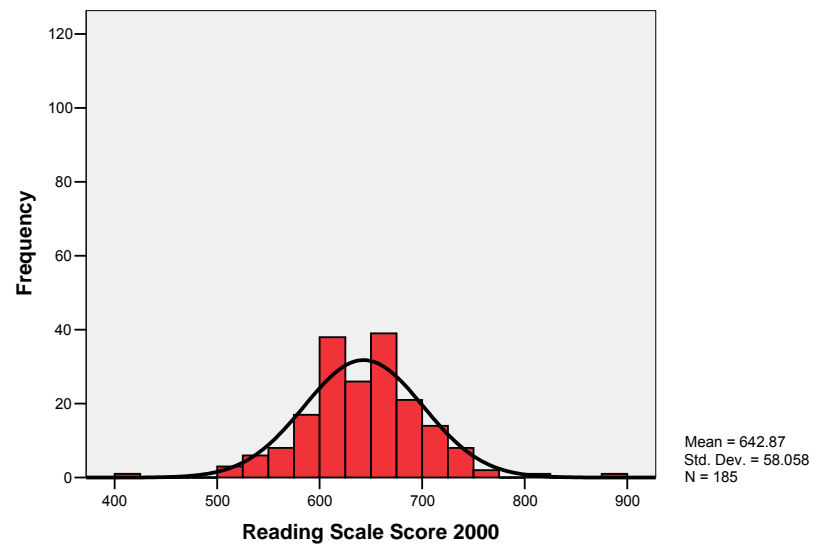


Figure 9a  
 Distribution of Reading Scores on the  
 Grade 8 CSAP, 2004 for  
 Non-IB Students

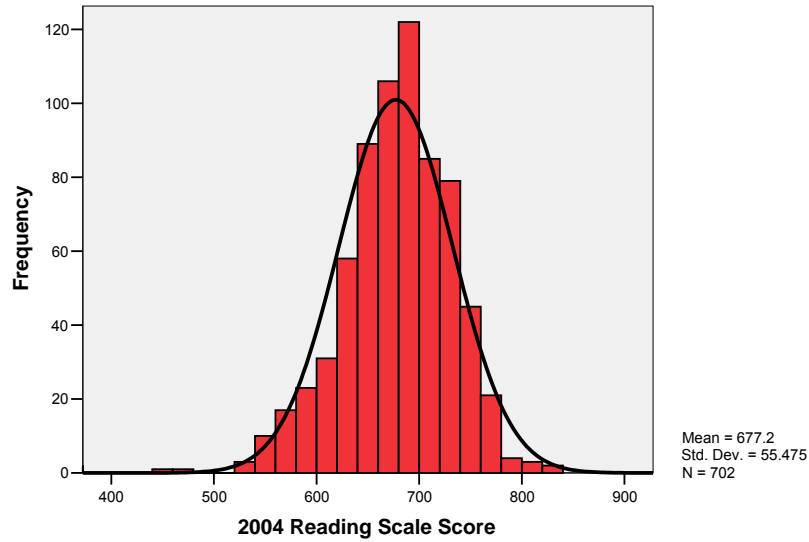
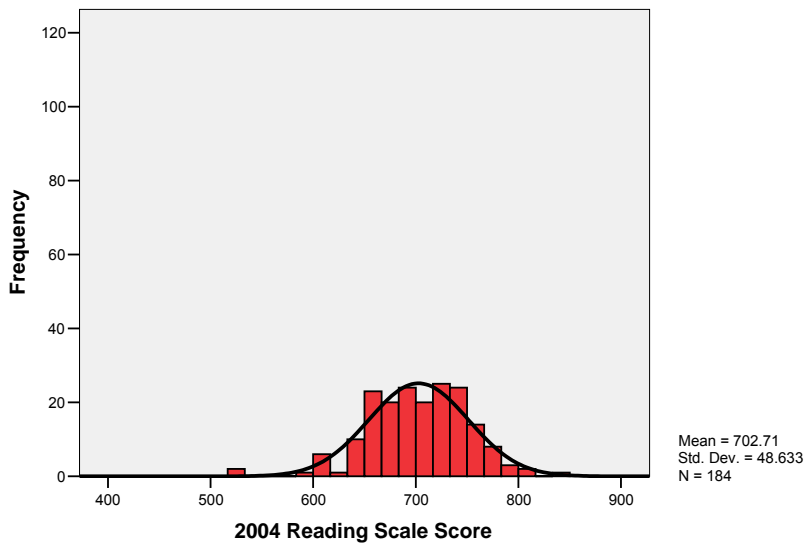


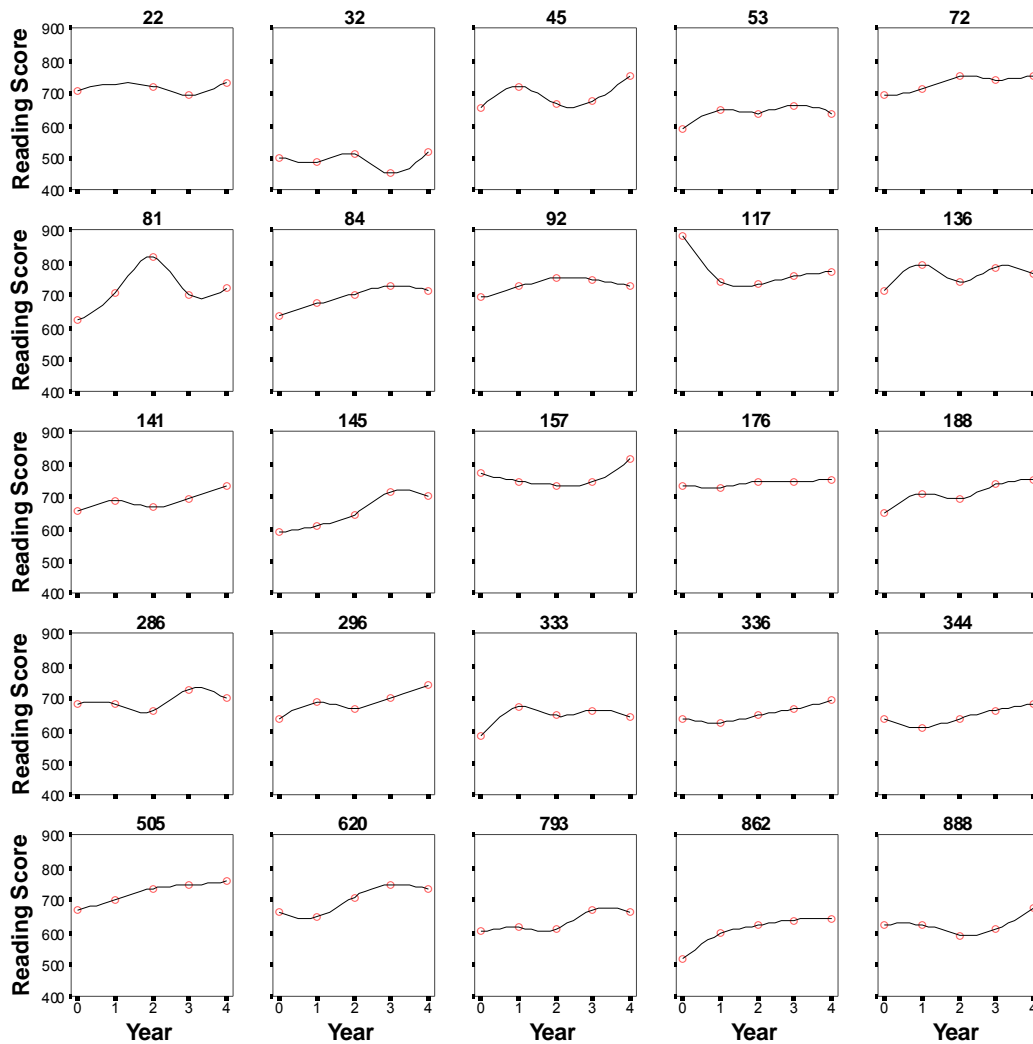
Figure 9b  
 Distribution of Reading Scores on the  
 Grade 8 CSAP, 2004 for  
 IB Students



In order to visualize how students' achievement change over time and to determine the best overall parametric model fit (e.g., linear, logistic) to the data we examine the “raw” empirical growth plots. Random samples of 25 IB and 25 non-IB students were selected and their empirical, nonparametric, growth plots examined.

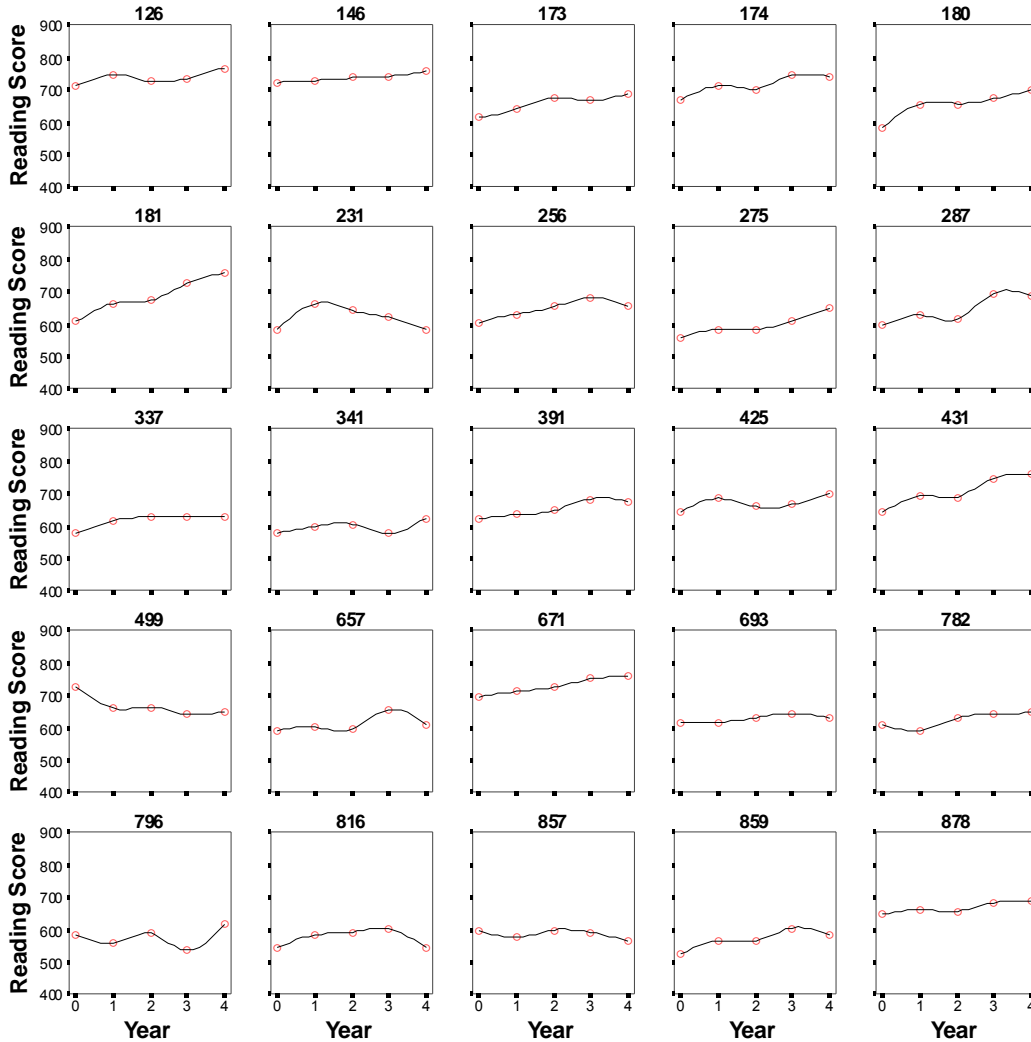
The “smoothed” empirical growth curves are illustrated in Figures 10a and 10b of IB and non-IB students, respectively.

Figure 10a  
 Nonparametric Reading Growth Curves  
 Grades 4-8, 2000-2004 for a  
 Sample of IB Students<sup>8</sup>



<sup>8</sup> The “id numbers” above each student plot are assigned at random and cannot be used to identify individual students.

Figure 10b  
 Nonparametric Reading Growth Curves  
 Grades 4-8, 2000-2004 for a  
 Sample of Non-IB Students<sup>9</sup>



In these two samples of 25 students it is apparent that the shapes of the reading growth curves are somewhat different for the two groups of students. That is, the empirical growth curves of the non-IB students are essentially linear, whereas those for the IB sample are more variable and exhibit more curvilinear forms. Also, there appears to be more variability in the intercepts and slopes of the IB students. Although several of the empirical growth trajectories in both samples are curvilinear, the best common functional form across individual trajectories appears to be linear.

<sup>9</sup> The “id numbers” above each student plot are assigned at random and cannot be used to identify individual students.

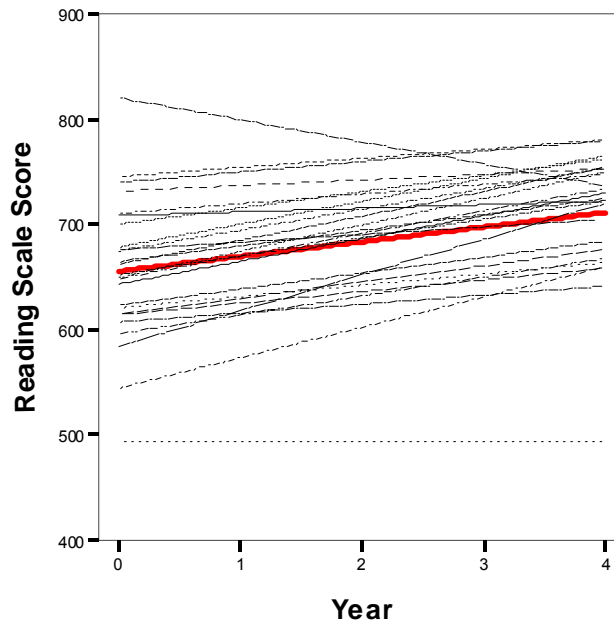
## 4.2.2 Results from the Multivariate Analyses

Results from the OLS fits and exploratory multiple regression predictive models are provided below.

### 4.2.2.1 Results from the Ordinary Least Squares Fits

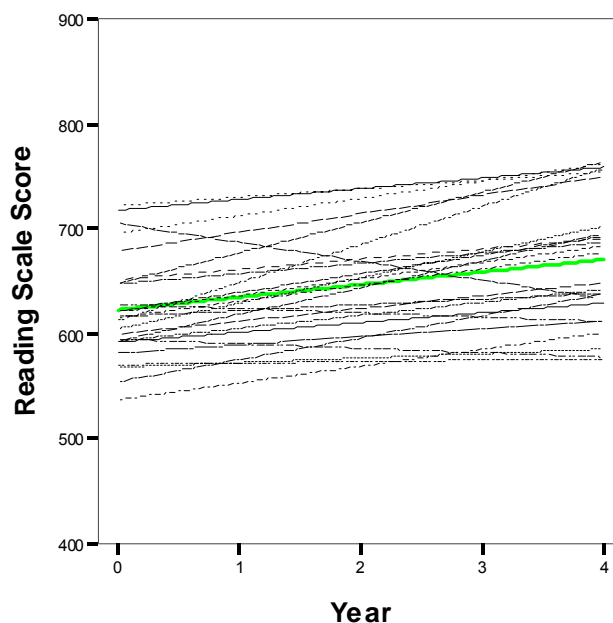
The ordinary least squares (OLS) fit to the raw achievement data indicates sufficient variability to warrant more sophisticated modeling. The OLS growth trajectories of the samples of 25 IB and 25 non-IB students are shown in Figures 11a and 11b, respectively.

Figure 11a  
Ordinary Least Squares Reading Growth Trajectories  
Grades 4-8, 2000-2004 for a Sample of IB Students



Legend: Black lines indicate individual growth trajectories.  
The red line is the average regression line for IB students.

Figure 11b  
 Ordinary Least Squares Reading Growth Trajectories  
 Grades 4-8, 2000-2004 for a Sample of Non-IB Students



Legend: Black lines indicate individual growth trajectories.  
 The green line is the average regression line for non-IB students.

Figures 11a-b indicate that there is substantial variability of the intercepts (2000 grade 4 test scores) for both IB and non-IB students. The intercepts of the IB students appear more variable than those non-IB students. The red line in Figure 11a and the green line in Figure 11b represent the average regression line for the two groups of students. Although the mean initial status (reading score in Year 0) appears higher in the IB sample than in the non-IB sample, the slopes of the regression lines (indicating growth rates) appear to be similar for these two groups of students. In addition, there appears to be less variability in the growth rates (slopes) than the intercepts, even though there are obvious outliers.

Figures 12a and 12b illustrate the mean regression line (center line) and the 95% confidence interval around the regression line (indicated by the two outer lines) for the samples of IB and non-IB students. The red circles represent individual reading scores.

The OLS regressions indicate that, on average, IB student scored almost 33 point higher on the grade 4 reading assessment (Year 0) and their rate of change (slope) was similar to that of non-IB students (14 vs. 12 points per year, respectively).

Figure 12a  
 Mean OLS Regression Line for Reading Growth  
 Grades 4-8, 2000-2004 for a Sample of IB Students

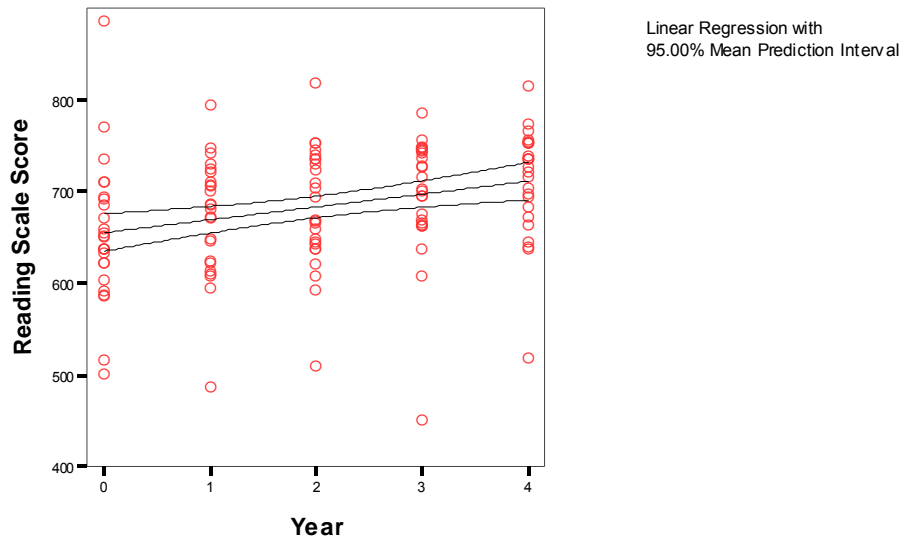
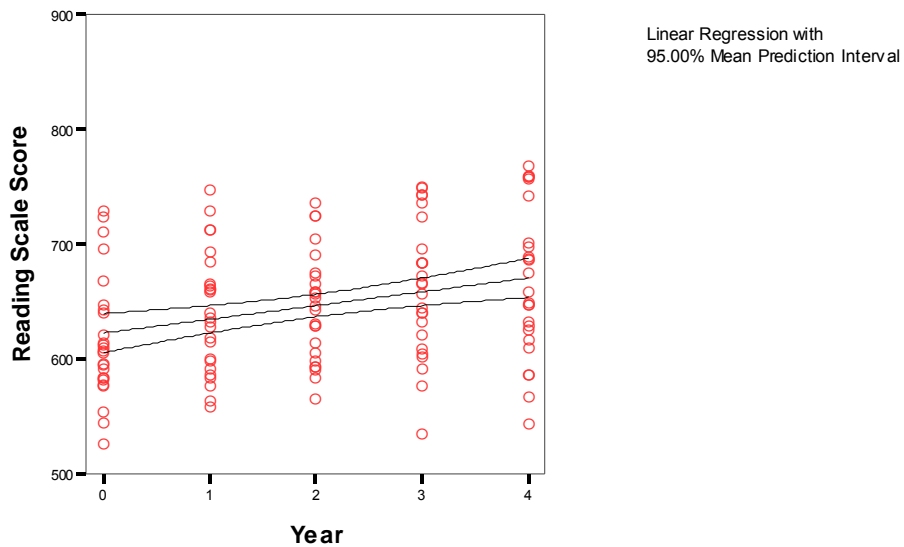


Figure 12b  
 Mean OLS Regression Line for Reading Growth  
 Grades 4-8, 2000-2004 for a Sample of Non-IB Students



4.2.2.2 Results from the Multiple Regression: Exploratory Predictive Models

In order to identify factors that are potentially predictive of student achievement and growth in reading, bivariate correlations and stepwise linear regression models are estimated prior to the hierarchical linear modeling. The bivariate correlations indicate that the strongest correlate to the outcome, reading achievement in grade 8, is initial status, i.e., reading score in fourth grade ( $r = .77$ ). The correlation between average annual growth in reading and reading outcomes

(grade 8) is much lower, but still statistically significant ( $r = .42$ ). In this cohort, there appears to be no correlation between initial reading status (grade 4) and average gain in reading from grade 4 through grade 8,  $r = .07$ .

Separate stepwise linear multiple regression models are specified in which the dependent (criterion) variables are (1) the Spring 2004 Grade 8 CSAP scores in reading and (2) average gain scores from the 2000 through 2004 CSAP reading assessments of Grades 4 through 8. The independent variables entered into the model as potential predictors are IB participation, total length of time in IB, prior achievement in reading (as measured by the 2000 CSAP), gender, minority status, poverty status, language background, cognitive disability, and non-cognitive disability.

*Outcomes models:* Stepwise regression allows the analysis program to select the variables that best predict the criterion variable, in this case score on the 2004 grade 8 reading assessment. The strongest variable selected by this procedure is initial status (grade 4 reading score), which accounts for approximately 60 percent of the variability in reading outcomes ( $R^2 = .599$ ). Although the results indicate that there is significant variation in reading outcomes still to be explained, none of the other variables, including IB participation, are adequate predictors of eighth grade reading performance.

In order to investigate the effect of IB participation<sup>10</sup>, two additional regression models are estimated. The first model entered IB status as the sole predictor of reading achievement in grade 8. This model indicates that IB participation has a small positive effect on reading achievement, but accounts for less than four percent of the variance ( $R^2 = .035$ ).

The second model estimates the effect of IB participation after controlling for the effect of initial status (reading performance in grade 4) on reading outcomes. After controlling for initial status, IB participation is no longer a significant predictor of reading achievement in the eighth grade. The coefficient (slope) for IB participation is not statistically significant. In other words, these results indicate that including IB in the equation does not explain any more of the variance than does initial reading status (2000 score) alone. This is evident by comparing the  $R^2$  for the model with initial status alone and the  $R^2$  for the model containing IB participation and initial status. The  $R^2$ s are identical, 0.599.

*Growth Models.* The regression analysis described above is an outcomes analysis since the criterion, or dependent, variable is eighth grade reading score (outcome) in 2004.

In order to provide a comparison between OLS results for reading *growth* and those obtained from a HLM repeated measures growth model a similar regression analysis with average reading growth as the criterion variable (rather than reading outcome in grade 8) was attempted. However, the correlations between average annual reading growth and IB participation and initial reading status are so low (.04 and .07, respectively) and statistically insignificant that the

---

<sup>10</sup> IB participation is defined here as 2 or more years in District 20's IB program during the period 2000 through 2004 (grades 4 through 8). Ninety-seven percent of IB students participated 2 or more years.

regression equations could not be estimated. Thus, these results indicate that neither IB nor initial status is related to growth in reading.

#### 4.2.3 Results from the HLM Estimation of the Effects of IB Participation on Reading Achievement and Growth in Middle School

The OLS fits to the raw achievement data discussed in § 4.2.2.1 indicate sufficient variability to warrant more sophisticated modeling. Hierarchical Linear Models (HLM) is used to examine the effects of IB participation on reading growth and outcomes.

##### 4.2.3.1 Results from the HLM Growth Models

The exploratory predictive multiple regression analyses indicate no significant effect of IB participation on either reading achievement or growth in middle school.

The results of the HLM analyses are consistent with these conclusions. A series of 2-level growth models are estimated using HLM. In multilevel models of individual change (i.e., growth modeling), Level 1 specifies the individual growth trajectory, which is modeled as dependent on person-level factors. These person-level effects are modeled in Level 2 as the slope coefficients for achievement growth. The dependent, or criterion, variable is student score on the grade 8 2004 CSAP reading assessment.

Initial preliminary unconditional models for means and growth indicate substantial variability in the intercepts (initial reading status) at Year 0 and slopes (growth rates) of students from grade 4 through grade 8 (2000-2004 grade 4-8 cohort). The first unconditional model is an unconditional *means* model, i.e., a model with no predictors in either level 1 or level 2. This model estimates the total amount of true variation in the outcome and allows us to initially partition the total variation in the outcome that is between and within persons. Initially partitioning the variance components provides an indication of whether there is enough variability within persons to warrant growth modeling and whether there is enough variability between persons to warrant a predictive model at level 2. The second unconditional model is an unconditional *growth* model, a model with time as the only level 1 predictor and no level 2 predictors. This type of model indicates whether we can account for the within-person variance by modeling growth alone or whether additional variability between persons in intercepts and slopes is great enough to model with level 2 predictors. The model equations and their results are presented below:

##### 4.2.3.2 Unconditional Means Model

The unconditional means model contains no predictors at either level and simply partitions the outcome variation rather than describing change over time. This unconditional model specifies that the growth trajectory for person  $i$  is completely flat, emanating from the intercept,  $\pi_{0i}$ , because the trajectory lacks the slope parameter associated with a temporal predictor.

**LEVEL 1 MODEL**

$$\text{READING}_{ti} = \pi_{0i} + e_{ti}$$

**LEVEL 2 MODEL**

$$\pi_{0i} = \beta_{00} + r_{0i}$$

Substituting the Level 2 coefficients into the Level 1 equation, we arrive at the combined model:

$$\text{READING}_{ti} = \beta_{00} + r_{0i} + e_{ti}$$

where

READING<sub>ti</sub> is the outcome variable, reading score in grade 8

π<sub>0i</sub> = True reading mean for person *i*

β<sub>00</sub> = Grand mean reading score across individuals and occasions

e<sub>ti</sub> = Level 1 residual variance in true intercept of person *i* (within-persons deviation)

r<sub>0i</sub> = Level 2 residual variance in true intercept across all individual in the population (between-persons deviation)

The model estimates are compared with those from the unconditional growth model, described below, in Table 9.

4.2.3.3 Unconditional Growth Model

The unconditional growth model introduces time (identified as Year) as the only level 1 predictor. No student predictors are specified in the Level 2 submodel. Based on results from the exploratory analyses described in § 4.2.1 and § 4.2.2, a linear change trajectory is specified.

**LEVEL 1 MODEL**

$$\text{READING}_{ti} = \pi_{0i} + \pi_{1i}(\text{YEAR}_{ti}) + e_{ti}$$

**LEVEL 2 MODEL**

$$\pi_{0i} = \beta_{00} + r_{0i}$$

$$\pi_{1i} = \beta_{10} + r_{1i}$$

Substituting the Level 2 coefficients into the Level 1 equation, we arrive at the combined model:

$$\text{READING}_{ti} = \beta_{00} + \beta_{10} * \text{YEAR}_{ti} + r_{0i} + r_{1i} * \text{YEAR}_{ti} + e_{ti}$$

where

READING<sub>ti</sub> is the outcome variable, reading score in grade 8

- $\pi_{0i}$  = Initial reading status of person  $i$ , that is, the expected outcome for that student in the spring of grade 4 (when Year = 0)
- $\pi_{1i}$  = Rate of change in reading (growth rate) for person  $i$
- YEAR is 0 at spring 2000, 1 at spring 2001, 2 at spring 2002, 3 at spring 2003 and 4 at Spring 2004 (grades 4, 5, 6, 7 and 8)
- $\beta_{00}$  = Estimated mean intercept, or initial reading status of a non-IB student at year 0
- $\beta_{10}$  = Mean academic year growth rate in reading
- $e_{it}$  = Level 1 residual variance in true growth trajectory of person  $i$  (within-persons deviation)
- $r_{0i}$  = Level 2 residual variance in true intercept across all individual in the population (between-persons deviation)
- $r_{1i}$  = Level 2 residual variance in true rate of change across all individual in the population (between-persons deviation)

The results of these two models are compared in Table 10

Table 10  
Comparison of Unconditional Means and Unconditional Growth Models

		Parameter	Unconditional Means Model	Unconditional Growth Model
<b>Fixed Effects</b>				
Initial status,		$\beta_{00}$	651.83*** (1.656)	621.34*** (1.716)
		$\pi_{0i}$		
Rate of change,		$\beta_{10}$		15.26*** (0.298)
		$\pi_{1i}$		
<b>Variance Components</b>				
Level 1	Within-persons	$e_{it}$	1194.82	546.46
Level 2	For initial status	$r_{0i}$	2207.20***	2301.48***
	For rate of change	$r_{1i}$		24.36***
	Covariance	$r_{01}$		- 15.30
<b>Goodness-of-fit</b>				
Deviance			45794.78	43364.73
$\chi^2$ statistic				2430.05***

Numbers in parentheses represent standard errors.

\*\*\*  $p < .001$

The intraclass correlation coefficient (ICC) is an estimate of the proportion of total variation in the outcome that lies “between” people, thus allowing us to compare the relative magnitude of the variance components. The ICC ( $\rho$ ) obtained from the unconditional means model is:

$$\rho = r_{0i} / (r_{0i} + e_{it})$$

$$\rho = 2207.20 / (2207.20 + 1194.82) = 0.6488$$

Thus, approximately 65 percent of the total variation in reading performance is attributable to differences between students. These results indicate that additional modeling is warranted.

Addition of a temporal component to the model (i.e., the slope of the growth trajectory) defines the intercept as the mean student score at the initial measurement (i.e., grade 4). This unconditional growth model indicates that students averaged about 621 ( $\beta_{00}$ ) on the grade 4 reading assessment and grew at an average rate of about 15 scale score points per year ( $\beta_{10}$ ) from fourth through eighth grades. The addition of the temporal variable significantly reduces the amount of residual (i.e., unexplained) variance at Level 1. However, a significant amount of between-persons residual variance in both initial status and rate of change remains to be explained (level 2). The unconditional growth model provides the baseline for modeling change over time. Comparing the variance components of subsequent models with the initial unconditional growth model provides an estimate of the amount of variation in outcome that is explained by the more complex models.

The next step is to model the effect of student characteristics and IB participation on both intercept (initial reading status) and rate of change (reading growth). Several conditional models are estimated. The first series of models examines the effect of IB participation (for 2 or more years) on reading achievement and growth from fourth through eighth grades. The second set of models evaluates the effect of *number of years of IB participation*. Both sets of models conceptualize IB as a “treatment” effect.

#### 4.2.3.4 HLM Growth Models: Effects of IB Participation on Reading Achievement and Growth

The “Exploratory Analysis” option of HLM is used to suggest the probable best Level 2 (student-level) predictors. Potential predictors include IB participation<sup>11</sup>, number of years in IB, gender, minority status, language background, and poverty status<sup>12</sup>. Two variables were suggested as potential predictors of the intercept (initial reading status): IB participation and years in the IB program. One variable was suggested as potentially effecting rate of change (the slope): initial reading status. None of the demographic variables appear to contribute to the explanation of reading achievement or growth in this cohort of students. This is not surprising since the descriptive analyses indicate few demographic differences between the IB and non-IB students.

Two models are discussed below. The first model evaluates the main effects of IB participation on both intercept and reading growth rate. The second model compares the effects of IB participation and initial status on growth. Thus, the second model is a value added model that at-

<sup>11</sup> IB participation is defined as student enrollment in the IB program for two or more years during the school years 1999-2000 through 2003-2004.

<sup>12</sup> Poverty status is determined by eligibility for free or reduced price lunch.

tempts to answer the question of whether the observed effect of IB participation (i.e., higher achievement) is due to selection effects, conceptualized here as prior reading achievement and operationalized as score on the 2000 grade 4 CSAP reading assessment.

#### 4.2.3.5 Main Effects of IB Participation

As described above, this model examines effects of IB participation on the intercept and slope without controlling for other student characteristics. The Level 1 and 2 submodels are:

##### **LEVEL 1 MODEL**

$$\text{READING}_{it} = \pi_{0i} + \pi_{1i} (\text{YEAR}_{it}) + e_{it}$$

##### **LEVEL 2 MODEL**

$$\pi_{0i} = \beta_{00} + \beta_{01} (\text{IB}) + r_{0i}$$

$$\pi_{1i} = \beta_{10} + \beta_{11} (\text{IB}) + r_{1i}$$

Substituting the Level 2 coefficients into the Level 1 equation, we arrive at the combined model:

$$\text{READING}_{it} = \beta_{00} + \beta_{01} (\text{IB}) + \beta_{10} (\text{YEAR}) + \beta_{11} (\text{IB} * \text{YEAR}) + r_{0i} + r_{1i} (\text{YEAR}_{it}) + e_{it}$$

where,

READING<sub>it</sub> is the outcome variable, reading score in grade 8

π<sub>0i</sub> = Initial reading status of person *i*, that is, the expected outcome for that student in the spring of grade 4 (when Year = 0)

π<sub>1i</sub> = Rate of change in reading (growth rate) for person *i*

YEAR is 0 at spring 2000, 1 at spring 2001, 2 at spring 2002, 3 at spring 2003 and 4 at Spring 2004 (grades 4, 5, 6, 7 and 8)

β<sub>00</sub> = Estimated mean intercept, or initial reading status of a non-IB student at year 0

β<sub>01</sub> = Mean IB/non-IB gap on the intercept (initial reading status)

β<sub>10</sub> = Mean academic year growth rate in reading

β<sub>11</sub> = Mean effect of IB on reading growth rate

IB = 0 if student was in IB program for less than 2 years or never in the program or 1 if student was in the IB program for 2 or more years

e<sub>it</sub> = Level 1 residual variance in true growth trajectory of person *i* (within-persons deviation)

r<sub>0i</sub> = Level 2 residual variance in true intercept across all individual in the population

r<sub>1i</sub> = Level 2 residual variance in true rate of change across all individual in the population (between-persons deviation)

The coefficients for the fixed effects and variance components are provided in Table 10 after discussion of the second model. On average, non-IB students in this cohort scored about 615 points (β<sub>00</sub>) on the 2000 grade 4 CSAP reading assessment, while IB students scored about 28 points higher (β<sub>01</sub>), averaging about 643 on the grade four assessment. Non-IB students aver-

aged a growth rate of 15.6 ( $\beta_{10}$ ) scale score points per year from fourth through eighth grades. The results suggest that IB students' average reading growth rate is slightly lower than that of non-IB students by 1.4 points ( $\beta_{11}$ ,  $p = .053$ ). That is, IB students' average reading growth rate is 14.2, as compared to 15.6 for non-IB students.

#### 4.2.3.6 Main Effects of IB Participation After Controlling for Initial Reading Status

As described above, the second conditional model examines effects of IB participation on the intercept and slope after controlling for selection effects as measured by score on the grade 4 CSAP reading assessment. This model is termed a "value-added model" because it evaluates the value added to reading growth by the IB program after accounting for the effects of initial status. A confounding factor is the fact that 62 percent of this cohort of IB students was enrolled in the IB program during the fourth grade and initial status is measured as reading score on the spring Grade 4 CSAP.

The Level 1 and 2 submodels are:

##### **LEVEL 1 MODEL**

$$\text{READING}_{it} = \pi_{0i} + \pi_{1i} (\text{YEAR}_{it}) + e_{it}$$

##### **LEVEL 2 MODEL**

$$\pi_{0i} = \beta_{00} + \beta_{01} (\text{IB}) + r_{0i}$$

$$\pi_{1i} = \beta_{10} + \beta_{11} (\text{IB}) + \beta_{12} (\text{READ2000}) + r_{1i}$$

Substituting the Level 2 coefficients into the Level 1 equation, we arrive at the combined model:

$$\text{READING}_{it} = \beta_{00} + \beta_{01} (\text{IB}) + \beta_{10} (\text{YEAR}) + \beta_{11} (\text{IB} * \text{YEAR}) + \beta_{12} (\text{READ2000}_i - \text{READ2000.}) * (\text{YEAR}_{it}) + r_{0i} + r_{1i} (\text{YEAR}_{it}) + e_{it}$$

where,

READING<sub>it</sub> is the outcome variable, reading score in grade 8

$\pi_{0i}$  = Initial reading status of person *i*, that is, the expected outcome for that student in the spring of grade 4 (when Year = 0)

$\pi_{1i}$  = Rate of change in reading (growth rate) for person *i*

YEAR is 0 at spring 2000, 1 at spring 2001, 2 at spring 2002, 3 at spring 2003 and 4 at Spring 2004 (grades 4, 5, 6, 7 and 8)

$\beta_{00}$  = Estimated mean intercept, or initial reading status of a non-IB student at year 0

$\beta_{01}$  = Mean IB/non-IB gap on the intercept (initial reading status)

$\beta_{10}$  = Mean academic year growth rate in reading

$\beta_{11}$  = Mean effect of IB on reading growth rate

$\beta_{12}$  = Mean effect of reading initial status on reading growth rate

IB = 0 if student was in IB program for less than 2 years or never in the program or 1 if student was in the IB program for 2 or more years

$e_{it}$  = Level 1 residual variance in true growth trajectory of person  $i$  (within-persons deviation)

$r_{0i}$  = Level 2 residual variance in true intercept across all individual in the population

$r_{1i}$  = Level 2 residual variance in true rate of change across all individual in the population (between-persons deviation)

The results for the IB main effects and value-added growth models are provided in Table 11.

Table 11  
Comparison of IB Main Effects Model and a Value Added Model <sup>a</sup>

		Parameter	IB Main Effects Growth Model	Value-Added Growth Model
<b>Fixed Effects</b>				
Initial status,		$\beta_{00}$	615.48*** (1.867)	615.48*** (1.867)
	$\pi_{0i}$	$\beta_{01}$	28.28*** (4.198)	28.28*** (4.198)
Rate of change,		$\beta_{10}$	15.55*** (0.335)	15.53*** (0.298)
	$\pi_{1i}$	$\beta_{11}$	- 1.41* (0.729)	- 1.31 <i>ns</i> (0.711)
		$B_{12}$		- 0.003 <i>ns</i> (0.007)
<b>Variance Components</b>				
Level 1	Within-persons	$e_{it}$	546.41	546.42
Level 2	For initial status	$r_{0i}$	2172.77***	2172.33***
	For rate of change	$r_{1i}$	24.15***	23.46***
	Covariance	$r_{01}$	- 8.88	- 0.43
Goodness-of-fit				
Deviance			43313.17	43323.66
$\chi^2$ statistic				3.75

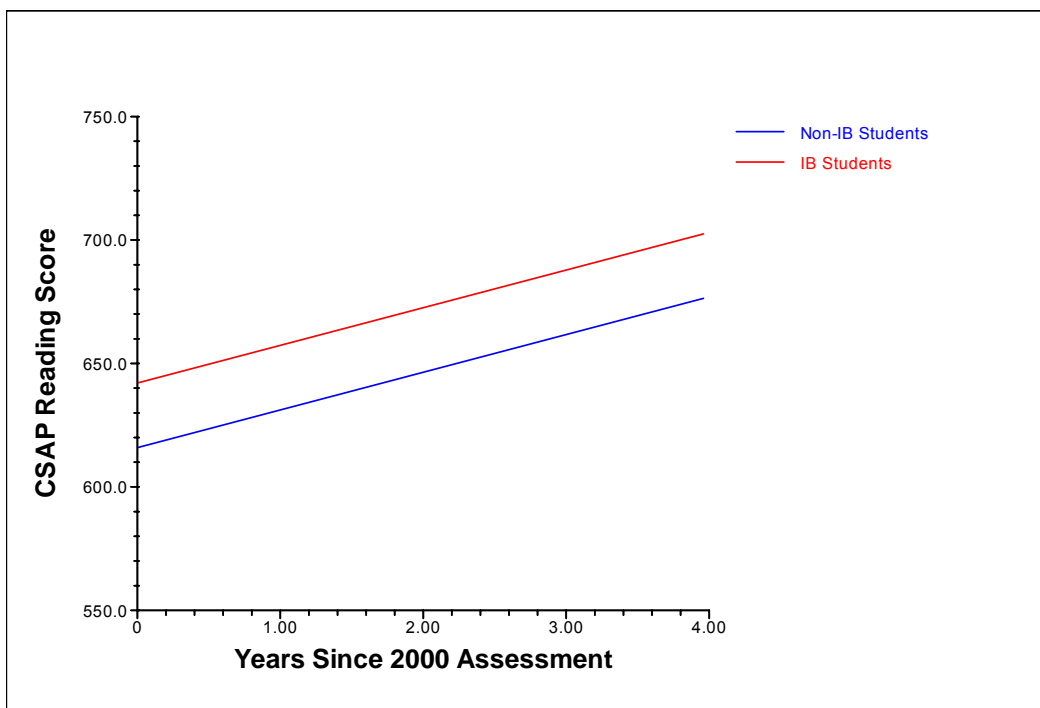
<sup>a</sup> This model evaluated the effect of IB on growth rates after accounting for initial status. Numbers in parentheses represent standard errors.

\*  $p \leq .05$  \*\*  $p \leq .01$  \*\*\*  $p < .001$  *ns* not statistically significant

As we saw in the previous IB main effects model, non-IB students, on average, score about 615 on the grade 4 reading assessment, while IB students score about 28 points higher. However,

when initial reading status is added as a predictor of reading growth, the significance level ( $p$  value) of IB participation on reading growth ( $\beta_{11}$ ) drops from .053 to .065, even though reading initial status does not appear to be related to growth rate ( $\beta_{11}$ ) in this cohort of students. These results suggest that the consistently higher performance of IB students may be attributable to selection effects and/or to the effects of other variables not available for these analyses rather than to IB participation. IB students do outperform other students on all CSAP assessments at all grade levels. However, if the IB program was having an additional, positive effect on student achievement, one would expect that the average reading growth trajectory of IB students would exhibit a steeper slope than that of non-IB students of similar ability. That is not the case, as indicated by Figure 13 below. This figure illustrates the average growth trajectories of IB and non-IB students. Note the parallel slopes of the two groups of students.

Figure 13  
Reading Growth Trajectories of IB and Non-IB Students  
Grades 4 – 8, 2000-2004



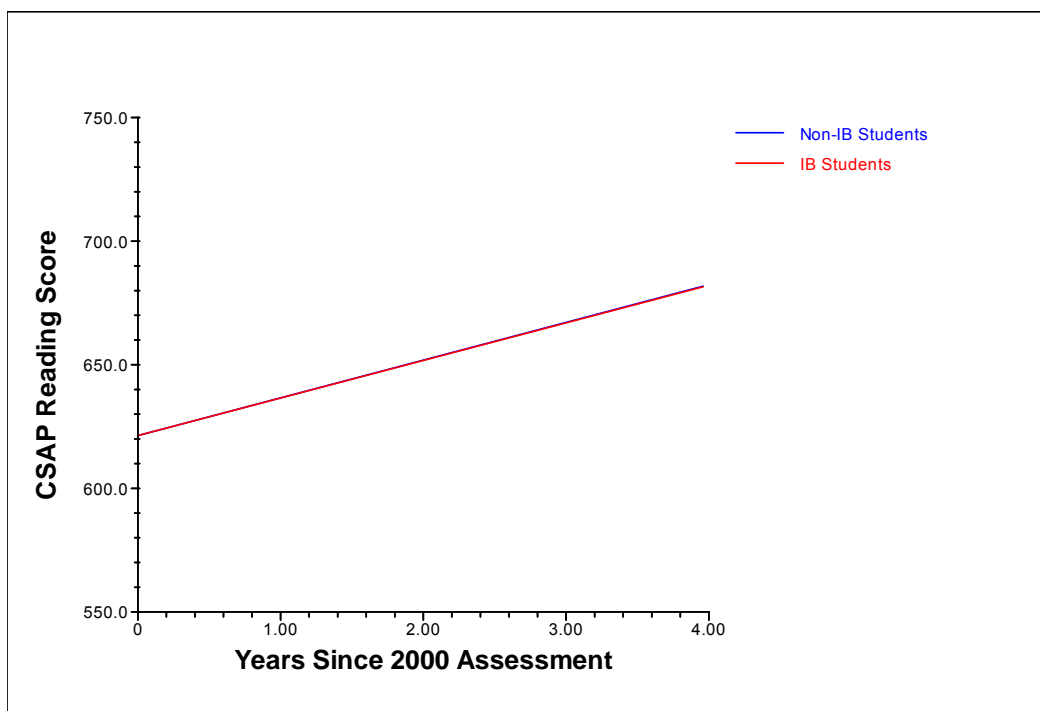
In order to examine possible differences between IB and non-IB students from a different viewpoint, a different type of growth model in which IB status is excluded from the model of the intercept (i.e., initial status) is estimated<sup>13</sup>. This model, described below, examines the effects of IB participation on reading growth rate without controlling for initial status or other student characteristics. If IB participation during elementary and middle school has a positive effect on reading achievement and growth, then we would expect to see a steeper slope for IB students among those who start at the same initial status, in this case the district average on the grade 4 reading

<sup>13</sup> Also, we obviously cannot attribute higher initial performance to subsequent IB participation. In effect, this is how the equations are specified in the HLM level 2 model presented on page 40.

assessment. In this model, the intercept is the grand mean for initial status (grade 4 reading mean score) across all students. The results are illustrated in Figure 14.

There appears to be only one line, but in reality, the red line exactly over-lays the blue line. This is because there is no difference in the slopes (growth rates) of the two groups of students. In other words, IB and non-IB students make, on average, exactly the same amount of gain in reading scores from fourth through eighth grades.

Figure 14  
Effect of IB Participation on Reading Growth Rate  
Grades 4-8, 2000-2004



The results of the Longitudinal Study of Student Growth (LSSG) demonstrated that in the 36 districts in the study, slope and initial reading status were inversely related. That is, students who start with high performance typically grow at a slower rate than lower performing students. That does not appear to be the case in this cohort of students. The correlation is essentially 0, indicating no relationship between reading performance and growth rate ( $r = -.04$ ). In order to rule out the possibility of a differential between the reading growth rates of very high scoring non-IB and IB students that might be attributable to the IB program, similar exploratory analyses were conducted for subgroups of students (non-IB and IB) who scored 600 or better on the grade 4 reading assessment or 650 or better on the grade 8 assessment. (In other words, these analyses examine whether the reading growth rates of the very high scoring IB students would likely be lower without participation in the program.) The results of those analyses essentially mirror the regression and HLM results presented above.

Number of years in the IB program was also suggested as a potential predictor of reading achievement and growth by HLM's Exploratory option. A 2-level model that evaluates the effect of number of years in the IB program on reading performance and growth was estimated. Goodness-of-fit statistics indicate that this model is no better than either the unconditional growth model or either of the two conditional models described above.

#### 4.2.4 Summary of Results for the 2000-2004 Grade 4-8 Cohort

In summary, these analyses indicate that the consistently higher performance of IB students in grades 4-8 may be attributable to selection effects and/or to the effects of other variables not available for these analyses. Some might argue that the most likely explanation for IB students' high performance is that they are brighter, more motivated, and more advantaged students. That is not the only possible explanation for the lack of definitive findings. The vast majority of students in District 20 are high achieving. Over 80 percent of this cohort of students scored proficient or advanced on the CSAP reading assessment each year. These students will be high achievers, regardless of what program they are in. Inarguably, IB students outscore non-IB students, but because of small numbers, small differences in variability around test scores and year-to-year growth, and general similarity between these two groups of students, there is little to discriminate between them.

### **4.3 Results for the 2001-2004 Grade 7-10 Cohort**

The results of the descriptive, OLS, multivariate, and HLM analyses of the 2001-2004 Grade 7-10 Cohort are described below.

#### 4.3.1 Description of the 2001-2004 Grade 7-10 Secondary School Cohort

There were 987 students in this cohort, 11 percent (107 students) of whom have been or are currently enrolled in the IB program during the secondary grades (7 through 10). As shown in Table 12, IB participation ranged from 1 year to 4 years during the period of the 2000-2001 through 2003-2004 school years. In this cohort, students tended to participate either two years or four years in the program, 36 percent and 53 percent, respectively.

Table 12  
 Number of Years of Participation in District 20's IB Program  
 2000-2004 Grade 4-8 Student Cohort

Years in IB Program	n	Percent
1	3	2.8
2	39	36.4
3	8	7.5
4	57	53.3
Total	107	100

In District 20, IB students in this secondary cohort are predominantly female (60 vs. 48 %). There are no other differences between IB and non-IB students with respect to minority status and language background.

Students in the IB program consistently out-perform students not in the program, as shown in Table 13. However, there is no significant difference between these subgroups of students in their rate of growth in reading from seventh through tenth grades.

Table 13  
 Mean Reading Scores<sup>a</sup> of IB and Non-IB Students<sup>b</sup>, Grades 7-8, 2001-2004

Reading Assessment	IB Students	Non-IB Students	Mean Difference	<i>p</i> value
2001 Grade 7	697	662	35	<.001
2002 Grade 8	714	675	39	<.001
2003 Grade 9	723	685	38	<.001
2004 Grade 10	742	701	41	<.001
Average Annual Gain, 2000-2004	15	13	2	.124

<sup>a</sup> Average scaled scores on the CSAP reading assessments.

<sup>b</sup> This cohort contains 107 IB students and 880 non-IB students.

The distributional characteristics of the IB/non-IB achievement indicate that while non-IB students span the range of scores on the CSAP, IB students are concentrated in the middle to upper ranges of CSAP, as illustrated in Figures 15a-b and 16a-b. Possible reading scores on the CSAP range from 300 to 900. The scores achieved by non-IB and IB students on the grade 7 CSAP spanned the range of 453 – 805 and 611 – 776, respectively. The ranges of scores on the grade 10 CSAP are 370 – 838 (non-IB) and 565 – 827 (IB).

Figure 15a  
Distribution of Reading Scores on the  
Grade 7 CSAP, 2001 for  
Non-IB Students

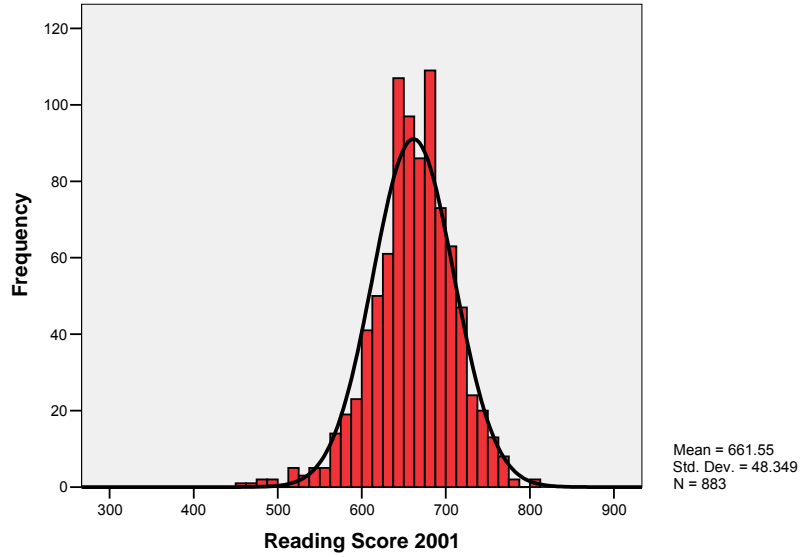


Figure 15b  
Distribution of Reading Scores on the  
Grade 7 CSAP, 2001 for  
IB Students

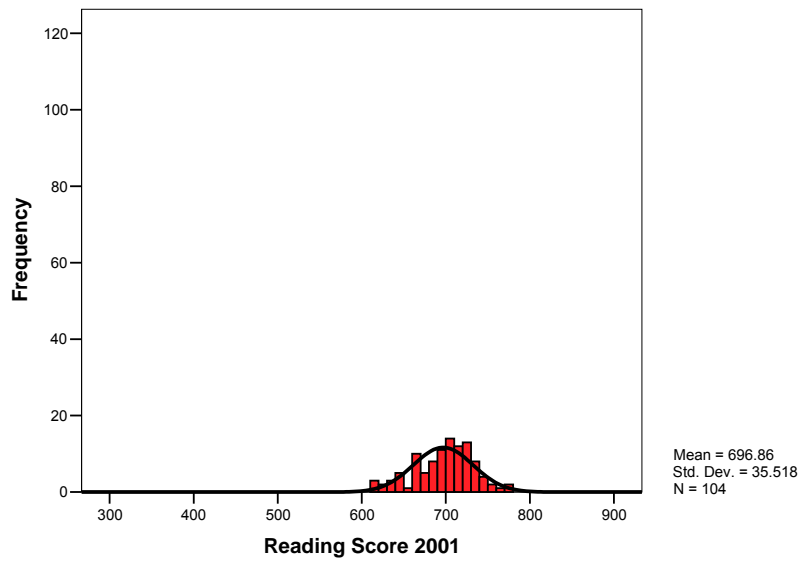


Figure 16a  
 Distribution of Reading Scores on the  
 Grade 10 CSAP, 2004 for  
 Non-IB Students

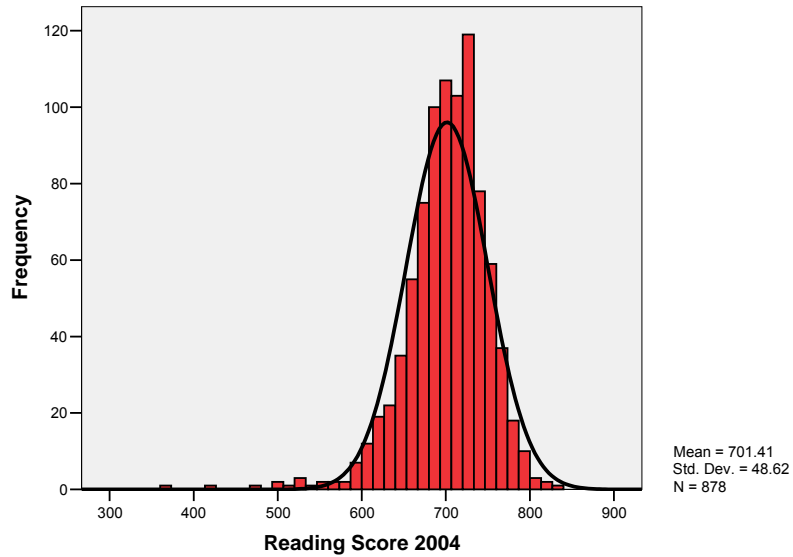
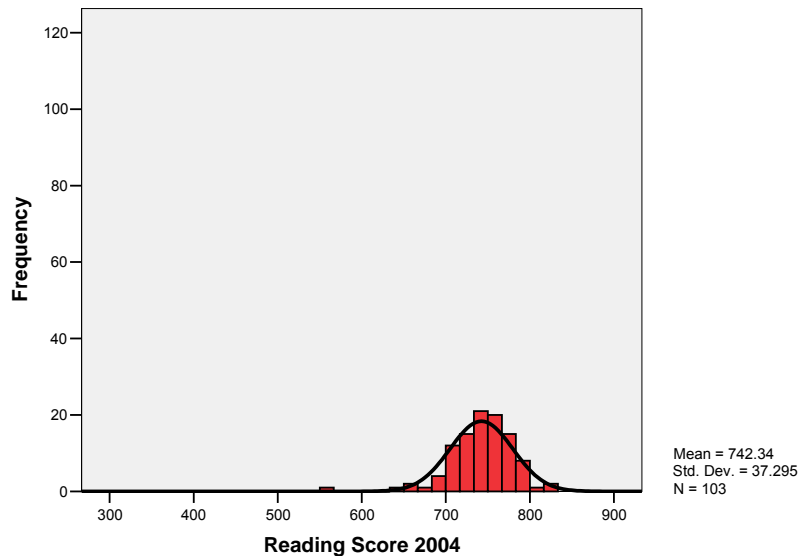
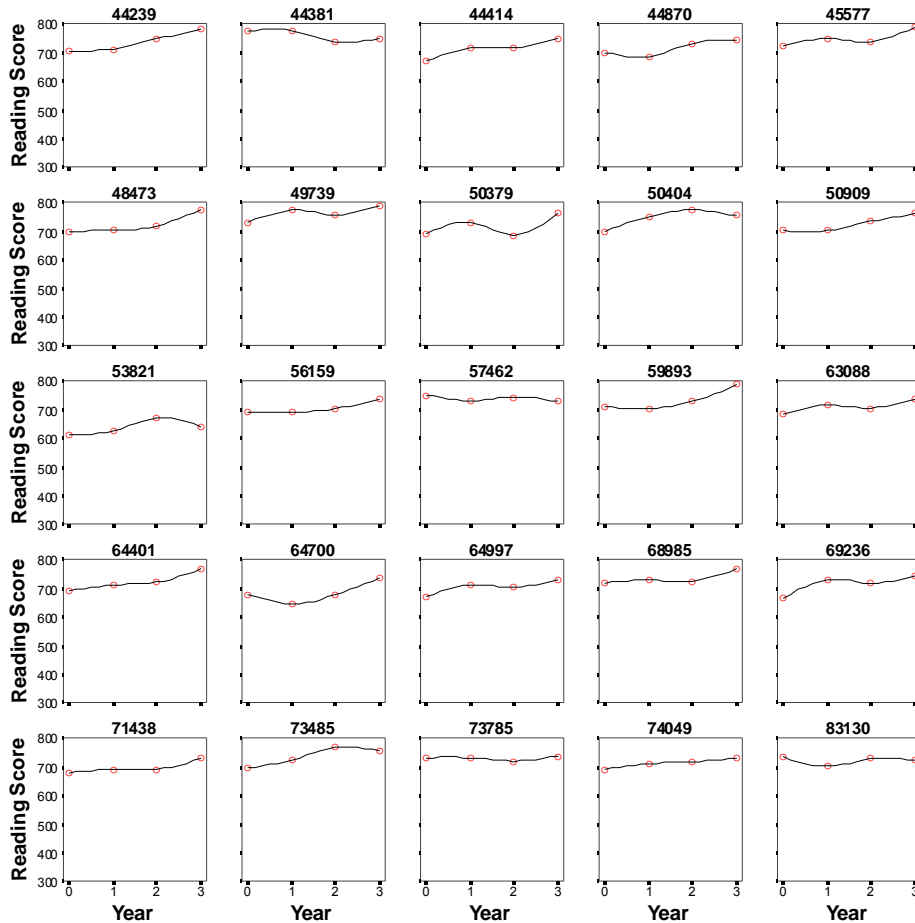


Figure 16b  
 Distribution of Reading Scores on the  
 Grade 10 CSAP, 2004 for  
 IB Students



The nonparametric growth plots of random samples of 25 IB and 25 non-IB students were examined. The “smoothed” empirical growth curves are illustrated in Figures 17a and 17b of IB and non-IB students, respectively.

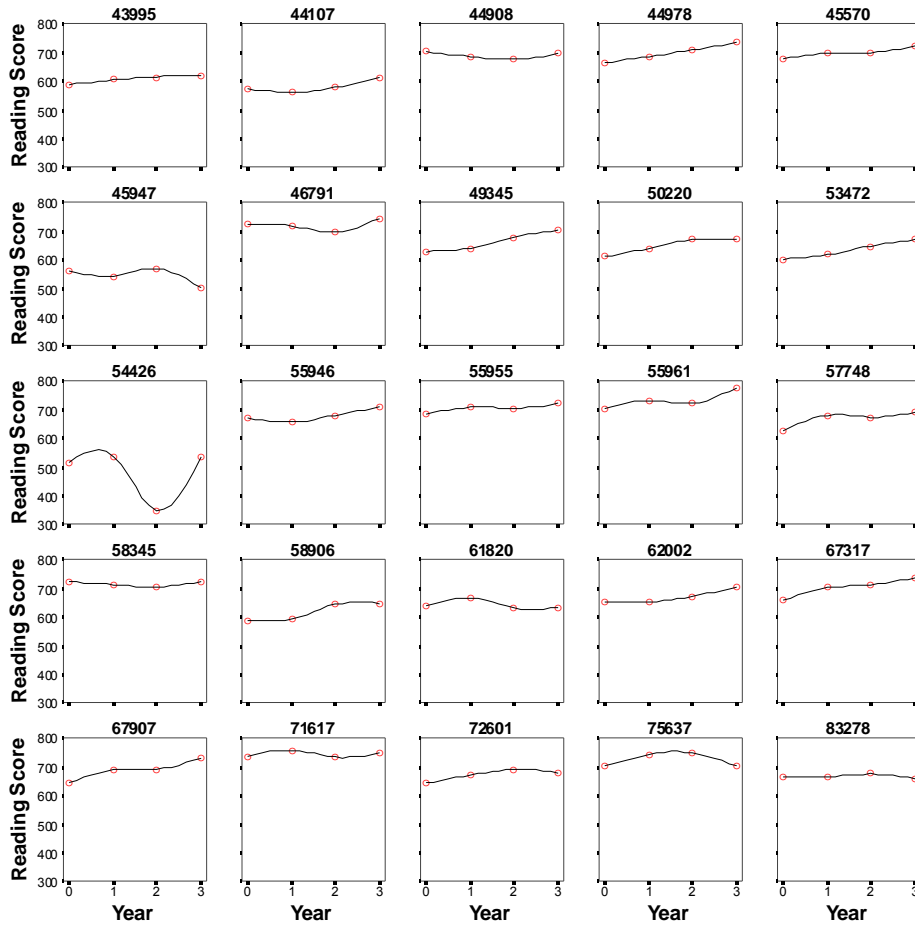
Figure 17a  
 Nonparametric Reading Growth Curves  
 Grades 7-10, 2001-2004 for a  
 Sample of IB Students<sup>14</sup>



In this sample of 50 students the shapes of the reading growth curves are similar for IB and non-IB students. The empirical growth curves are essentially linear and positive, with a few notable exceptions (students 45947, 54426 in Figure 17b). Also, there appears to be more variability in the intercepts of the non-IB students. Although several of the empirical growth trajectories in both samples are curvilinear, the best common functional form across individual trajectories appears to be linear.

<sup>14</sup> The “id numbers” above each student plot are assigned at random and cannot be used to identify individual students.

Figure 17b  
 Nonparametric Reading Growth Curves  
 Grades 7-10, 2001-2004 for a  
 Sample of Non-IB Students<sup>15</sup>



### 4.3.2 Results from the Multivariate Analyses

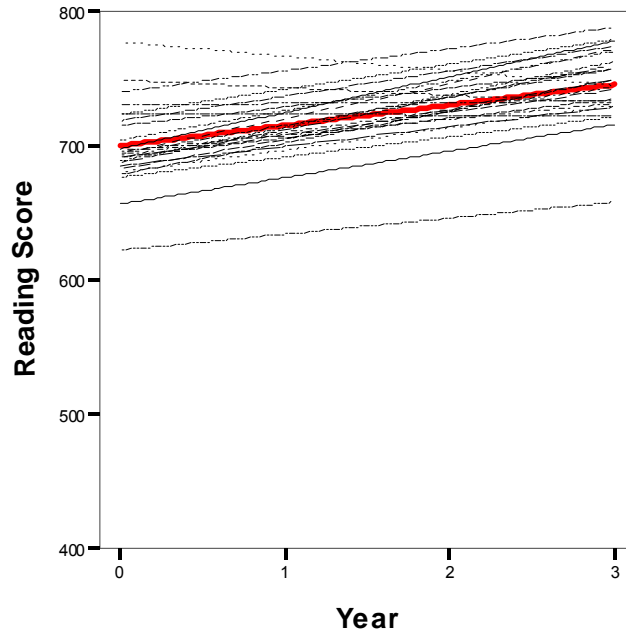
Results from the OLS fits and from the exploratory multiple regression predictive models are discussed in the sections below.

#### 4.3.2.1 Results from the Ordinary Least Squares Fits

The ordinary least squares (OLS) fit to the raw achievement data indicates sufficient variability to warrant more sophisticated modeling. The OLS growth trajectories of the samples of 25 IB and 25 non-IB students are shown in Figures 18a and 18b, respectively.

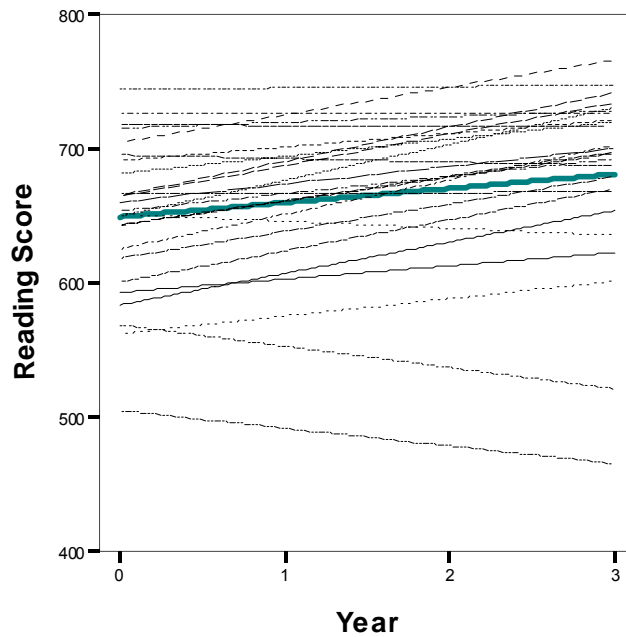
<sup>15</sup> The “id numbers” above each student plot are assigned at random and cannot be used to identify individual students.

Figure 18a  
 Ordinary Least Squares Reading Growth Trajectories  
 Grades 7-10, 2001-2004 for a Sample of IB Students



Legend: Black lines indicate individual growth trajectories.  
 The red line is the average regression line for IB students.

Figure 18b  
 Ordinary Least Squares Reading Growth Trajectories  
 Grades 7-10, 2001-2004 for a Sample of Non-IB Students



Legend: Black lines indicate individual growth trajectories.  
 The green line is the average regression line for non-IB students.

Figures 18a and 18b indicate that there is substantially more variability in the intercepts (2001 grade 7 test scores) and slopes of non-IB students than the IB students. The red line in Figure 18a and the green line in Figure 18b represent the average regression lines for the two groups of students. Although the mean initial status (reading score in Year 0) is higher in the IB sample than in the non-IB sample, the mean regression slopes (indicating growth rates) appear to be similar for these two groups. In addition, there appears to be less variability in the growth rates (slopes) than the intercepts, even though there are obvious outliers.

Figures 19a and 19b illustrate the mean regression line (center line) and the 95% confidence interval around the regression line (indicated by the two outer lines) for the samples of IB and non-IB students. The red circles represent individual reading scores. These plots illustrate dramatically the greater variability in the reading scores of the non-IB students during middle and high school.

The OLS regressions indicate that, on average, IB students scored 35 points higher on the grade 7 reading assessment (Year 0), and their rate of change (slope) was similar to that of non-IB students (15 vs. 13 points per year, respectively).

Figure 19a  
 Mean OLS Regression Line for Reading Growth  
 Grades 7-10, 2001-2004 for a Sample of IB Students

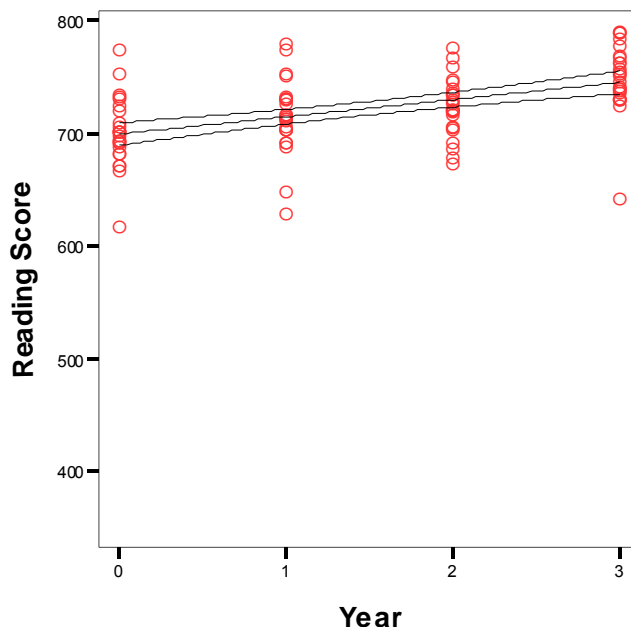
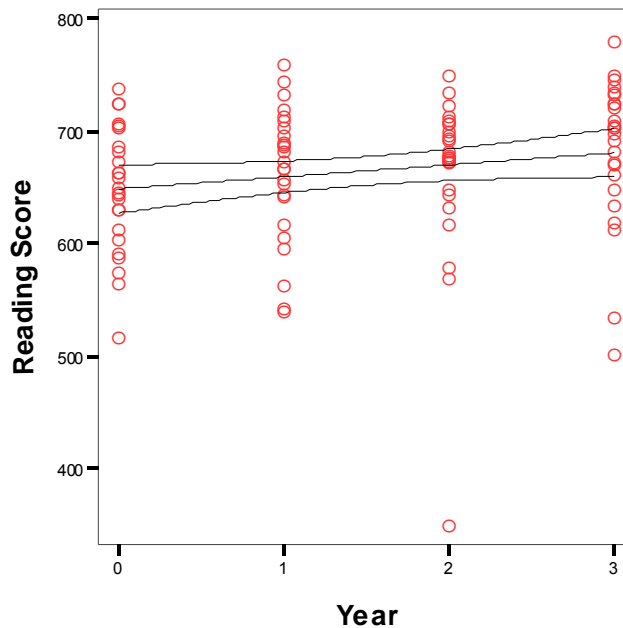


Figure 19b  
 Mean OLS Regression Line for Reading Growth  
 Grades 7-10, 2001-2004 for a Sample of Non-IB Students



#### 4.3.2.2 Multiple Regression: Exploratory Predictive Models

In order to identify factors that are potentially predictive of student achievement and growth in reading, bivariate correlations and stepwise linear regression models are estimated prior to the hierarchical linear modeling. The bivariate correlations indicate that the strongest correlate to the outcome, reading achievement in grade 10, is initial status, i.e., reading score in grade 7 ( $r = .732$ ). The correlation between average annual growth in reading and initial status (grade 7) is much lower and negative,  $r = -.338$ . On the other hand, the correlation between reading growth rate and reading outcomes (grade 10) is positive,  $r = .402$ . However, the correlation between average annual reading growth and IB participation is very low and not statistically significant ( $r = .047$ ).

Stepwise linear multiple regression models are specified in which the dependent (criterion) variables are the Spring 2004 Grade 10 CSAP scores in reading and average gain scores from the 2001 through 2004 CSAP reading assessments of Grades 7 through 10. The independent variables entered into the model as potential predictors are IB participation, length of time in IB, prior achievement in reading (as measured by the 2001 CSAP), gender, poverty status, cognitive disability, and non-cognitive disability.

*Outcomes Models:* Stepwise regression allows the analysis program to select the variables that best predict the criterion variable, in this case score on the 2004 grade 10 reading assessment. The strongest variable selected by this procedure is initial status (grade 7 reading score), which accounts for approximately 52 percent of the variability in reading outcomes on the 2004 grade

10 reading assessment ( $R^2 = .525$ ). Although the results indicate that there is significant variation in reading outcomes still to be explained, none of the other variables, including IB participation, are adequate predictors of tenth grade reading performance.

In order to investigate the effect of IB participation<sup>16</sup>, two additional regression models are estimated. The first model enters IB status as the sole predictor of reading achievement in grade 10. This model indicates that IB participation has a small positive effect on reading achievement, but that most of the variability remains to be explained.  $R^2 = .065$ , indicating that only about six percent of the variance in tenth grade reading achievement is explained by participation in the IB program.

The second model estimates the effect of IB participation after controlling for the effect of initial status (reading performance in grade 7) on reading outcomes. After controlling for initial status, IB participation retains a very small but statistically significant effect on reading achievement in the tenth grade. The change in  $R^2$  is .009, indicating that IB participation increases the amount of variance explained in tenth grade reading performance by about 1 percent. The total amount of variation explained by both initial status and IB participation is 53.9 percent. The facts that initial status alone accounts for 53 percent of the variance, IB participation alone accounts for 6.5 percent of the variance, but together the two predictors together account for only 53.9 percent indicates that most of the variance attributed to IB participation may be attributable to initial status. However, it also could be argued that some of the variance attributed to initial status is, in fact, attributable to prior IB participation (i.e., in elementary and early middle school) which cannot be evaluated here to lack of data for the earlier years.

*Growth Models.* The regression analyses described above are *outcomes* analyses since the criterion, or dependent, variable is tenth grade reading score (outcome) in 2004. The results discussed below are from a multiple regression growth model, since the dependent variable is average reading gain from 2001 through 2004.

The stepwise multiple regression, in which the strongest predictor of the criterion variable, average reading gain from 2001-2004, is entered first, followed by the next best predictor, and so on, indicates that initial status (reading score on the grade 7 assessment) is the best predictor of observed reading growth, accounting for 11.9 percent of the variance ( $R^2 = .119$ ). Thus, about 12 percent of the variance in growth in reading scores is due to initial reading status; however, a significant portion of the variance, 88 percent, remains to be explained. The next variable entered into the predictive equation is IB participation, which explains an additional 1.6 percent of the residual variance (total  $R^2 = .135$ ). Gender accounts for another 1.4 percent of the variance (total  $R^2 = .149$ ). Thus, the three variables, initial status, IB participation and gender, explain only 15 percent of the variance in reading growth, leaving 85 percent still to be explained. Together, the results indicate that initial status may be the best, albeit an insufficient, predictor of growth in reading achievement from grade 7 through 10 and that IB participation has an almost imperceptible effect on reading growth.

---

<sup>16</sup> IB participation is defined here as 2 or more years in District 20's IB program during the period 2001 through 2004 (grades 7 through 10). Ninety-seven percent of IB students participated 2 or more years.

### 4.3.3 Results from the HLM Estimation of Achievement Growth

The results of the HLM analyses are consistent with these conclusions. A series of 2-level growth models are estimated using Hierarchical Linear Models (HLM). The dependent variable is student score on the grade 10 2004 CSAP reading assessment.

Preliminary unconditional models for means and growth indicate substantial variability in the intercepts (initial reading status) at Year 0 and slopes (growth rates) of students from grade 7 through grade 10. The first unconditional model is an unconditional *means* model, i.e., a model with no predictors in either level 1 or level 2. This model estimates the total amount of true variation in the outcome and allows us to initially partition the total variation in the outcome that is between and within persons. Initially partitioning the variance components provides an indication of whether there is enough variability within persons to warrant growth modeling and whether there is enough variability between persons to warrant a predictive model at level 2. The second unconditional model is an unconditional *growth* model, a model with time as the only level 1 predictor and no level 2 predictors. This type of model indicates whether we can account for the within-person variance by modeling growth or whether additional variability between persons in intercepts and slopes is great enough to model with level 2 predictors. The model equations and their results are presented below:

#### 4.3.3.1 Unconditional Means Model

The unconditional means model contains no predictors at either level and simply partitions the outcome variation rather than describing change over time. This unconditional model specifies that the growth trajectory for person  $i$  is completely flat, emanating from the intercept,  $\pi_{0i}$ , because the trajectory lacks the slope parameter associated with a temporal predictor (for change over time). Thus, the estimated outcome for a student is the person's reading mean plus measurement error.

##### **LEVEL 1 MODEL**

$$\text{READING}_{ti} = \pi_{0i} + e_{ti}$$

##### **LEVEL 2 MODEL**

$$\pi_{0i} = \beta_{00} + r_{0i}$$

Substituting the Level 2 coefficients into the Level 1 equation, we arrive at the combined model:

$$\text{READING}_{ti} = \beta_{00} + r_{0i} + e_{ti}$$

where

$\text{READING}_{ti}$  is the outcome variable, reading score in grade 10

$\pi_{0i}$  = True reading mean for person  $i$

$\beta_{00}$  = Grand mean reading score across individuals and occasions

$e_{ti}$  = Level 1 residual variance in true intercept of person  $i$  (within-persons deviation)

$r_{0i}$  = Level 2 residual variance in true intercept across all individual in the population  
(between-persons deviation)

The model estimates are compared with those from the unconditional growth model, described below, in Table 3.

#### 4.3.3.2 Unconditional Growth Model

The unconditional growth model introduces time as the only level 1 predictor. No student predictors are specified in the Level 2 submodel. Based on the exploratory analyses described in § 4.3.1 and § 4.3.2, a linear change trajectory is specified.

##### **LEVEL 1 MODEL**

$$\text{READING}_{ti} = \pi_{0i} + \pi_{1i}(\text{YEAR}_{ti}) + e_{ti}$$

##### **LEVEL 2 MODEL**

$$\pi_{0i} = \beta_{00} + r_{0i}$$

$$\pi_{1i} = \beta_{10} + r_{1i}$$

Substituting the Level 2 coefficients into the Level 1 equation, we arrive at the combined model:

$$\text{READING}_{ti} = \beta_{00} + \beta_{10} * \text{YEAR}_{ti} + r_{0i} + r_{1i} * \text{YEAR}_{ti} + e$$

where

$\text{READING}_{ti}$  is the outcome variable, reading score in grade 10

$\pi_{0i}$  = Initial reading status of person  $i$ , that is, the expected outcome for that student in the spring of grade 7 (when Year = 0)

$\pi_{1i}$  = Rate of change in reading (growth rate) for person  $i$

YEAR is 0 at spring 2001, 1 at spring 2002, 2 at spring 2003 and 3 at Spring 2004 (grades 7, 8, 9, and 10)

$\beta_{00}$  = Estimated mean intercept, or initial reading status of a non-IB student at year 0

$\beta_{10}$  = Mean academic year growth rate in reading

$e_{ti}$  = Level 1 residual variance in true growth trajectory of person  $i$  (within-persons deviation)

$r_{0i}$  = Level 2 residual variance in true intercept across all individual in the population (between-persons deviation)

$r_{1i}$  = Level 2 residual variance in true rate of change across all individual in the population (between-persons deviation)

The results of these two models are provided in Table 14.

Table 14  
Comparison of Unconditional Means and Unconditional Growth Models  
2001 – 2004

		Parameter	Unconditional Means Model	Unconditional Growth Model
<b>Fixed Effects</b>				
Initial status,		$\beta_{00}$	684.36***	664.82***
	$\pi_{0i}$		(1.421)	(1.515)
Rate of change,		$\beta_{10}$		13.04***
	$\pi_{1i}$			(0.373)
<b>Variance Components</b>				
Level 1	Within- persons	$e_{it}$	768.18	381.97
Level 2	For initial status	$r_{0i}$	1799.31***	1998.51***
	For rate of change	$r_{1i}$		60.40***
	Covariance	$r_{01}$		- 80.29
	Correlation	$r_{01}$		- 0.231
Goodness-of-fit				
Deviance			39406.24	37931.96
$\chi^2$ statistic				1474.28***

Numbers in parentheses represent standard errors.

\*\*\*  $p < .001$

The unconditional means model stipulates that an individual's *true* growth trajectory is completely flat, and originates from the intercept,  $\pi_{0i}$  (person *i*'s average reading scores regardless of time). This model, which does not include a temporal predictor (slope), simply partitions the total variation in the outcome into within person and between person components. The intra-class correlation coefficient (ICC) is an estimate of the proportion of total variation in the outcome that lies "between" people, thus allowing us to compare the relative magnitude of the variance components. The ICC ( $\rho$ ) obtained from the unconditional means model is:

$$\rho = r_{0i} / (r_{0i} + e_{it})$$

$$\rho = 1799.31 / (1799.31 + 768.18) = 0.7008$$

Thus, approximately 70 percent of the total variation in reading performance is attributable to differences between students. These results indicate that additional modeling is warranted.

When a temporal component is included in the model (i.e., the slope of the growth trajectory), the intercept represents student scores at the initial measurement (i.e., grade 7). Thus, students average about 665 on the initial grade 7 reading assessment ( $\beta_{00}$ ) and grow at an average rate of about 13 scale score points per year ( $\beta_{10}$ ) from seventh through tenth grades. The addition of the temporal variable (unconditional growth model) significantly reduces the amount of residual (i.e., unexplained) variance at Level 1 by 50 percent  $[(768.18 - 381.97) / 768.18] = .503$ . However, a significant amount of between-persons residual variance in both initial status and rate of change remains to be explained (level 2). The unconditional growth model provides the baseline for modeling change over time. Comparing the variance components of subsequent models with the initial unconditional growth model provides an estimate of the amount of variation in outcome that is explained by the more complex model.

The next step is to model the effect of student characteristics and IB participation on both intercept (initial reading status) and rate of change (reading growth). Several conditional models are estimated. The initial conditional models indicate that none of the hypothesized covariates (e.g., initial status, gender) significantly predict reading growth once IB variables are accounted for. The two models described below examine the effect of IB participation (for 2 or more years) on reading achievement and growth from seventh through tenth grades and the effect of *number of years of IB participation*, respectively. Both models conceptualize IB as a “treatment” effect.

#### 4.3.3.3 HLM Growth Models: Effects of IB Participation on Reading Achievement and Growth

The “Exploratory Analysis” option of HLM is used to suggest the probable best Level 2 (student-level) predictors. Potential predictors include IB participation<sup>17</sup>, number of years in IB, initial reading status, gender, and poverty status<sup>18</sup>. Three variables were suggested as potential predictors of rate of change (the slope): IB participation, years in the IB program, and initial reading status. None of the demographic variables appear to contribute to the explanation of reading achievement or growth in this cohort of students. This is not surprising since the descriptive analyses indicate few demographic differences between the IB and non-IB students.

Two models are discussed below. The first model evaluates the main effect of IB participation on reading growth rate. The second model compares the effects of number of years in the secondary IB program on reading growth. In order to estimate a value-added model that attempts to answer the question of whether the observed effect of IB participation is due to selection effects, reading growth rate is modeled as a function of initial status in addition to the IB effects. There does not appear to be a significant effect of initial reading status on reading growth rate, while the main effects of both IB participation and length of time in the secondary IB program are significant. However, the results regarding prior achievement (operationalized here as reading achievement in grade 7) are inconclusive and should not be over-interpreted because student data are not available prior to grade 7 in 2001. We do not know students’ (1) reading

---

<sup>17</sup> IB participation is defined as student enrollment in the IB program for two or more years during the school years 1999-2000 through 2003-2004.

<sup>18</sup> Poverty status is determined by eligibility for free or reduced price lunch.

achievement, (2) IB participation, or (3) the effect of IB participation on reading achievement prior to grade 7. We do know, however, that IB students' "initial status" in grade 7 is confounded by their prior exposure to the IB program. Thus, the true effect of initial status (prior to IB exposure) cannot be estimated.

#### 4.3.3.4 Main Effect of IB Participation

While students who were enrolled in IB for at least two years from grades 7 through 10 score significantly higher on the grade 7 reading assessment than students who were not (697 vs. 661), obviously we cannot attribute higher initial performance to subsequent IB participation. However, the higher initial status of IB students (in 7<sup>th</sup> grade) can be attributed to selection effects, including prior performance and IB history. We simply do not know how the effects of prior participation in the IB program are manifested.

Because of this difficulty in interpretation of IB effects on initial status and implications for interpretation (or misinterpretation) of growth rates, a model that eliminates subsequent IB effects on initial status is estimated. This model examines effects of IB participation on reading growth rate without controlling for other student characteristics. The Level 1 and 2 submodels are:

#### **LEVEL 1 MODEL**

$$\text{READING}_{it} = \pi_{0i} + \pi_{1i} (\text{YEAR}_{it}) + e_{it}$$

#### **LEVEL 2 MODEL**

$$\pi_{0i} = \beta_{00} + r_{0i}$$

$$\pi_{1i} = \beta_{10} + \beta_{11} (\text{IB}) + r_{1i}$$

Substituting the Level 2 coefficients into the Level 1 equation, we arrive at the combined model:

$$\text{READING}_{it} = \beta_{00} + \beta_{10} (\text{YEAR}) + \beta_{11} (\text{IB} * \text{YEAR})) + r_{0i} + r_{1i} (\text{YEAR}_{it}) + e_{it}$$

where,

READING<sub>it</sub> is the outcome variable, reading score in grade 10

π<sub>0i</sub> = Initial reading status of person *i*, that is, the expected outcome for that student in the spring of grade 7 (when Year = 0)

π<sub>1i</sub> = Rate of change in reading (growth rate) for person *i*

YEAR is 0 at spring 2001, 1 at spring 2002, 2 at spring 2003 and 3 at Spring 2004 (grades 7, 8, 9 and 10)

β<sub>00</sub> = Estimated mean intercept, or initial reading status of students at year 0

β<sub>10</sub> = Mean academic year growth rate in reading

β<sub>11</sub> = Mean effect of IB on reading growth rate

IB = 0 if student was in IB program for less than 2 years or never in the program or 1 if

student was in the IB program for 2 or more years (during 7<sup>th</sup>-10<sup>th</sup> grades)  
 $e_{it}$  = Level 1 residual variance in true growth trajectory of person  $i$  (within-persons deviation)  
 $r_{0i}$  = Level 2 residual variance in true intercept across all individual in the population  
 $r_{1i}$  = Level 2 residual variance in true rate of change across all individual in the population  
 (between-persons deviation)

The results from this model are shown in Table 15.

Table 15  
 Effects of IB Participation on Reading Growth in Grades 7 – 10  
 2001 – 2004

		Parameter	Coefficient
<b>Fixed Effects</b>			
Initial status,		$\beta_{00}$	664.82*** (1.515)
$\pi_{0i}$			
Rate of change,		$\beta_{10}$	12.52*** (0.394)
$\pi_{1i}$		$\beta_{11}$	4.95*** (1.046)
<b>Variance Components</b>			
Level 1	Within- persons	$e_{it}$	381.86
Level 2	For initial status	$r_{0i}$	1999.17***
	For rate of change	$r_{1i}$	61.52***
	Covariance	$r_{01}$	-97.66
	Correlation	$r_{01}$	- 0.278
Goodness-of-fit			
Deviance			37910.04

Numbers in parentheses represent standard errors.

\*\*\*  $p < .001$

In this model, the intercept ( $\beta_{00}$ ) is the grand mean for initial status (grade 7 reading mean score) across all students. The coefficient,  $\beta_{11}$ , represents the effect of IB participation on reading growth rate ( $\beta_{10}$ ). Students in this cohort scored, on average, about 665 points on the 2001 grade 7 CSAP reading assessment ( $\beta_{00}$ ) and averaged a growth rate of about 13 scale score points per year ( $\beta_{10}$ ) from seventh through tenth grades. The results suggest that IB students' reading growth rate is slightly higher than the average growth rate by 5 points ( $\beta_{11}$ ,  $p < .001$ ).

That is, IB students' average reading growth rate is about 18 points, as compared to the overall average of 13 points. The variance decomposition is also provided in Table 15. Unfortunately, it appears that the addition of IB participation explains very little of the residual between-person variance in rate of change ( $r_{i1} = 60.4$  for the unconditional growth model, and  $r_{i1} = 61.5$  in the conditional model).

Figure 20 represents a different look at the effect of IB participation on reading growth rates from seventh through tenth grades. The plot depicts the average growth trajectories of IB and non-IB students. The intercept is the district average of 665 scale score points on the grade 7 reading assessment in 2001. The growth trajectories illustrate the small, but significant, effect of IB participation during late middle and early high school on reading achievement and growth. Starting from the district average, IB students appear to have a 5 point advantage in growth rate.

The more common approach is to illustrate the growth trajectories of IB and non-IB students, projected from their respective average intercepts, as seen in Figure 21. Figure 21 depicts the estimated growth trajectories of students who score at the average of their groups, IB and non-IB students. The value in considering the information illustrated by Figure 20 is that controlling for starting point, IB students appear to progress from grade 7 through 10 at a slightly faster rate than non-IB students.

Figure 20  
 Effect of IB Participation on Reading Growth Rate  
 Grades 7-10, 2001-2004

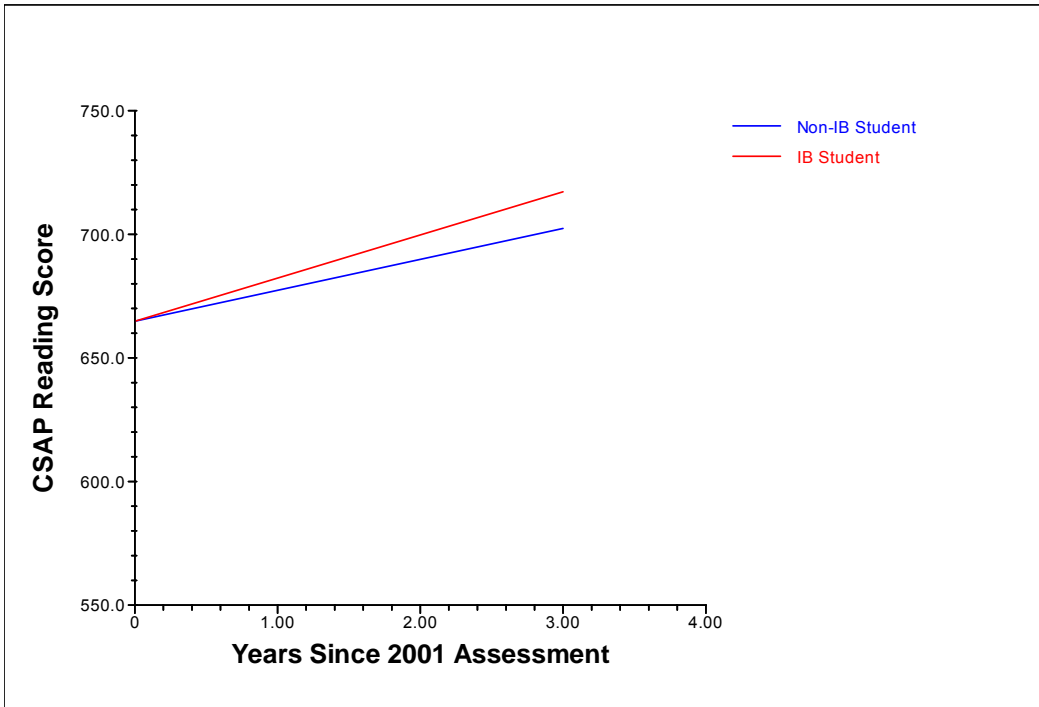
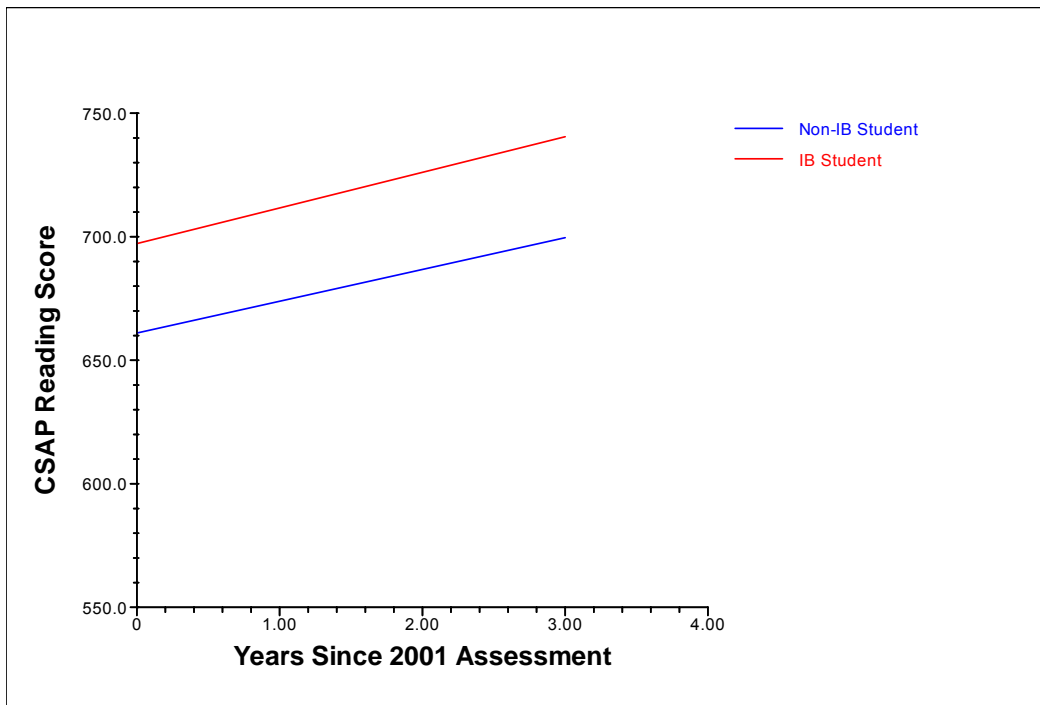


Figure 21  
 Reading Growth Trajectories of IB and Non-IB Students  
 Grades 7 – 10, 2001-2004



#### 4.3.3.5 Main Effect of Length of Time in the Secondary IB Program (Grades 7 – 10)

As described above, the second conditional model examines effect of length of participation in the secondary IB program on the grade 10 CSAP reading assessment and on the reading growth rate from grade 7 through grade 10. The previous model indicates that IB participation may have a small positive effect on reading growth. This may be termed a “treatment effect.” This begs the next question: Is length of time in the secondary IB program related to reading growth rate? In other words, given that IB participation has a positive influence on reading growth rate, do the benefits of the program accrue over time? Do students who spend more time in the program derive greater benefit? The model described below provides an estimate of the amount of reading growth that may be attributable to each additional year of participation in the IB program from grade 7 to 10.

The Level 1 and 2 submodels are:

##### **LEVEL 1 MODEL**

$$\text{READING}_{it} = \pi_{0i} + \pi_{1i} (\text{YEAR}_{it}) + e_{it}$$

##### **LEVEL 2 MODEL**

$$\pi_{0i} = \beta_{00} + r_{0i}$$

$$\pi_{1i} = \beta_{10} + \beta_{11} (\text{Time in IB}) + r_{1i}$$

Substituting the Level 2 coefficients into the Level 1 equation, we arrive at the combined model:

$$\text{READING}_{it} = \beta_{00} + \beta_{10} (\text{YEAR}) + \beta_{11} (\text{Time in IB} * \text{YEAR}) + r_{0i} + r_{1i} (\text{YEAR}_{it}) + e_{it}$$

where,

$\text{READING}_{it}$  is the outcome variable, reading score in grade 8

$\pi_{0i}$  = Initial reading status of person  $i$ , that is, the expected outcome for that student in the spring of grade 7 (when Year = 0)

$\pi_{1i}$  = Rate of change in reading (growth rate) for person  $i$

YEAR is 0 at spring 2001, 1 at spring 2002, 2 at spring 2003 and 3 at Spring 2004 (grades 7, 8, 9 and 10)

$\beta_{00}$  = Estimated mean intercept, or initial reading status of students at year 0

$\beta_{10}$  = Mean academic year growth rate in reading

$\beta_{11}$  = Mean effect of length of time in IB on reading growth rate

Time in IB = 0 through 4 (0 years in IB, 1 year in IB, 2 years in IB, etc., during grades 7 - 10)

$e_{it}$  = Level 1 residual variance in true growth trajectory of person  $i$  (within-persons deviation)

$r_{0i}$  = Level 2 residual variance in true intercept across all individual in the population  
 $r_{1i}$  = Level 2 residual variance in true rate of change across all individual in the population  
 (between-persons deviation)

Table 16  
 Effects of Length of IB Participation on Reading Growth during Grades 7 – 10  
 2001 – 2004

		Parameter	Coefficient
<b>Fixed Effects</b>			
Initial status,		$\beta_{00}$	664.82*** (1.515)
$\pi_{0i}$			
Rate of change,		$\beta_{10}$	12.51*** (0.393)
$\pi_{1i}$		$\beta_{11}$	1.594*** (0.273)
<b>Variance Components</b>			
Level 1	Within- persons	$e_{it}$	382.04
Level 2	For initial status	$r_{0i}$	1998.75***
	For rate of change	$r_{1i}$	61.39***
	Covariance	$r_{01}$	-99.46
	Correlation	$r_{01}$	- 0.284
Goodness-of-fit			
Deviance			37909.77

Numbers in parentheses represent standard errors.

\*\*\*  $p < .001$

The previous model indicates that IB participation in itself may have a positive effect on reading growth. The model described in Table 16 estimates the amount of reading growth that may be attributable to each additional year of participation in the IB program from grade 7 to 10. As shown previously, the district average in initial status (grade 7 reading score) is 665 ( $\beta_{00}$ ), with an average growth rate of about 13 points per year ( $\beta_{10}$ ). However, for each year of participation in IB during these grades, the growth rate increases by about one and a half scale score points per year ( $\beta_{11}$ ), for an approximate additional boost of nearly six and one-half points for the student who participates all four years. Thus, the predicted 2004 grade 10 reading score for a non-IB student scoring at the district average is 704:

$$665 + (13 * 3) = 704.$$

In comparison, the additional boost from IB during grades 7 – 10 is about 1.6 points per year of participation, so that the predicted grade 10 reading score for IB participants would be 710.4 for 4-year participation:

$$665 + (13 * 3) + (1.6 * 2) = 707.2 \text{ for two years of participation,}$$

$$665 + (13 * 3) + (1.6 * 3) = 708.8 \text{ for three years of participation, and}$$

$$665 + (13 * 3) + (1.6 * 4) = 710.4 \text{ for four years of participation.}$$

These analyses indicate that the total accrued effect of IB participation for four years of participation is a total of six points on the CSAP reading assessment. Although statistically significant, the total accrued effect appears to be less than one percent.

## 5. Conclusions

All analyses indicate that IB students consistently out-perform non-IB students in reading. While analyses of the middle school cohort (i.e., grades 4-8) indicate that this higher performance may be attributable to selection effects and/or to the effects of other variables not available for analysis, analyses of the youngest and oldest cohorts (grades 3-5 and 7-10) indicate that for students scoring at the same level in grade 7, IB may have a slightly positive effect on reading growth rate during the middle and early high school years. While some might argue that the most likely explanation for IB students' high performance is that they are brighter and more motivated, the results presented above also indicate the possibility of an accrual of benefits from participation in the IB program over time. The analysis of the effect of length of time in the IB program during 7<sup>th</sup> through 10<sup>th</sup> grades indicates that for each year of participation in IB, reading growth rate increases by about one and a half scale score points per year, for an approximate additional boost of six and one-half points for the student who participates all four years. To put this apparent increase in perspective, however, the reader must remember that, although statistically significant, the total accrual effect appears to be less than one percent.

The lack of definitive evidence of positive effects of IB participation on reading achievement and growth should *not* be viewed as negative evidence. The homogeneity and typically high performance of all three cohorts of students make it difficult to discriminate between groups of students and evaluate the effectiveness of many school programs on achievement. Add to that (a) the inability to measure many variables hypothesized to be related to achievement, (b) self-selection and program effectiveness (including achievement prior to the measurements included in this study, IB participation during earlier grades, and family SES), and (c) the inefficiency of the proxy variables available for analysis, estimation of true program effects is not as clear nor as precise as we would wish.