Academy School
District Twenty

Effects of IB Participation on Mathematics Achievement and Growth, 2001-2004

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

October 7, 2005

Executive Summary

This report describes the results of the longitudinal study of the effects of the International Baccalaureate (IB) program on mathematics 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 solely 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 mathematics.

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 most striking results from these analyses are:

- Although average grade 5-8 mathematics scores are higher for students in the middle years IB program, their growth rate appears lower than for non-IB students. Thus, the middle school program does not appear to provide an obvious acceleration to students' learning growth in mathematics.
- Conversely, IB students in the diploma program demonstrate significant average annual gains over their non-IB counterparts.
- Students in the IB diploma program score significantly higher on the CSAP mathematics assessments of grades 8 through 10.
- Additional years of participation in the IB program during middle and high school appear to be associated with greater gains in mathematics achievement in high school.
- Students' mathematics performance in grade 5 is strongly related to their reading ability in grade 5.
- Students' mathematics performance in grade 8 is strongly related to their reading ability in grade 8.
- Unlike reading performance and growth, growth in math during middle or high school does not appear to be related to prior achievement.
- Males out-score females by 15 points, on average, on the grade 8 mathematics assessment.

Although this study suggests that IB participation has a positive impact on student achievement and gains in mathematics during middle and high school, the main effects are relatively small. The homogeneity and typically high performance of District 20 students makes it difficult to discriminate between groups of students and to evaluate the effectiveness of many school programs designed to increase achievement. Another likely explanation is interaction between participation in the IB program and selection effects (e.g., students' academic ability and prior achievement, interest, motivation, family background, parental involvement, etc.) that cannot be disentangled. We also cannot know how these IB students would have performed without the IB program.

Table of Contents

Effects of IB Participation on Mathematics Achievement and Growth, 2001-2004

1. Introduction

This report is the second in a series of reports that describes the results of a multi-component study to examine the effects of IB program participation on student achievement and growth in student achievement in reading and mathematics at the elementary and secondary levels. The first report was entitled *Effects of IB Participation on Reading Achievement and Growth, 2000-2004 in Academy School District Twenty*. This report describes the results of the longitudinal study of the effects of the International Baccalaureate (IB) program on mathematics achievement and growth from 2001 through 2004 in grades 5-8, and 8-10. The secondary study cohorts are overlapping in grades, but not individuals, since students in grades 5 through 8 during the period 2001-2004 and in grades 8 through 10 during 2002-2004 are examined.

This report tells the "math story" in Academy School District Twenty from 2001 through 2004. Described are the primary research questions, methodology, and results of the study of the effects of participation in the IB program in secondary schools on student achievement and growth in mathematics achievement (as measured by CSAP) from 2001-2004. 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 solely 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 mathematics.

2. Research Questions

This phase of the research project investigates four 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 mathematics achievement and growth?
2. Do the effects of IB vary by school level (grades 5-8 and 8-10)?
3. Is there an interaction between student characteristics and the effects of the IB program or mathematics 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 mathematics achievement and growth?

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).

For the purposes of this study, the analytic cohorts (2001-2004 grades 5-8 and 2002-2004 grades 8-10) are defined as those students who had valid scores on the 2001/2002 and 2004 CSAP math assessments. HLM permits missing test data in computing the growth trajectories, but requires complete data on all predictors. Since initial math status (scores on the 2001 grade 5 and 2002 grade 8 assessments) is a predictor of math outcomes and growth[1], only students who had complete data on these assessments could be included in the analytical cohorts. Depending on the analytical cohort and type of analysis (outcomes or growth) reading initial status and reading outcomes (scores on the 2001 grade 5, 2002 grade 8, and 2004 grades 8 and 10 reading assessments) are also predictors of math outcomes and growth; thus, only students who also had complete data on these assessments could be included in those particular analyses.

For analytic purposes, middle school students who had participated in the program for two or more years during 6-8 grade are classified as "IB students." Students who never participated in IB or for only one year in middle school are classified as "non-IB." High school students who were in the IB program in the 2003-2004 school year (i.e., during tenth grade) are classified as "IB students;" all others are classified as "non-IB." Preliminary analyses indicate that there are significant performance differences between students who were in the IB program in tenth grade and students in the program in ninth grade (2002-2003), but who dropped out (n = 6) of the program by tenth grade. Thus, the total number of students and the number of students in these cohorts will not match the numbers of students in these grades each year or the numbers of students in the IB programs.

3.1 Descriptive Analyses

Prior to conducting multivariate analyses and hierarchical modeling, descriptive statistics are examined. Frequency distributions, crosstabulation tables, and t-tests (for differences in means) compare IB students and non-IB students. The variables compared include CSAP math and reading scores, 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.

[1] Much prior research has shown that the subject area performance having the highest correlation with mathematics achievement is reading. Therefore, CSAP reading scores are included as predictors in the models of mathematics performance.

3.2 Multiple Regression and Correlation

In order to identify factors that are potentially predictive of student achievement and growth in mathematics, bivariate correlations and stepwise multiple linear regressions are completed prior to estimating the hierarchical linear models. In addition, multiple linear regression is the analytic procedure used to model achievement outcomes, i.e., to determine the effects of IB participation, prior achievement in mathematics and reading, and student background on mathematics achievement.

The dependent (criterion) variables are the Spring 2004 CSAP scores in mathematics and average gain scores from the 2001 or 2002 through 2004 CSAP mathematics 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[2]:

Table 1
Analytical Cohorts by Year and Grade

Cohort	Year/Grade			
	2001	2002	2003	2004
2002-2004 Grade 5-8 Cohort[3]	5	6	7	8
2000-2004 Grade 8-10 Cohort		8	9	10

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	
2001-2004 Grade 5-8 Cohort	2004 Grade 8 CSAP Math Score	Avg. Gain in Math 2001-2004
2002-2004 Grade 8-10 Cohort	2004 Grade 10 CSAP Math Score	Avg. Gain in Math 2002-2004

Potential independent variables (predictors) include IB participation, length of time in IB, prior achievement in mathematics (as measured by the initial CSAP assessment listed above), reading achievement (as measured by CSAP), gender, minority status, poverty status, and language background. The possible interaction of gender and IB participation also are 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

[2] Year refers to the Spring CSAP administrations.
[3] Cohort years refer to the times of the Spring CSAP administrations. Thus, 2001 refers to the 2000-2001 school year.

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 mathematics 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 on student growth in mathematics achievement, controlling for student background characteristics. The HLM models estimated are models of individual change. Estimation of the effects on individual change over time in the multilevel context is known colloquially as growth modeling.

The HLM 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.

The dependent variable in the growth models is student score on the 2004 CSAP mathematics 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 mathematics, reading achievement, gender, minority, poverty status, language background, cognitive disability, and IB status (the “treatment effect”).

Prior to estimation of IB effects on mathematics 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 model specifies the growth trajectory and level 2 models the effects of individual factors. The model in this example examines effects of IB participation on mathematics growth rate without controlling for other student characteristics. The Level 1 and 2 submodels are:

LEVEL 1 MODEL

$$MATH_{ti} = \pi_{0i} + \pi_{1i} (YEAR_{ti}) + e_{ti}$$

LEVEL 2 MODEL

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

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

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

$$MATH_{ti} = \beta_{00} + \beta_{10} (YEAR) + \beta_{11} (IB*YEAR)) + r_{0i} + r_{1i} (YEAR_{ti}) + e_{ti}$$

where,

$MATH_{ti}$ is the outcome variable, mathematics score in 2004
π_{0i} = Initial mathematics status of person *i*, that is, the expected outcome for that student in the spring of Year = 0
π_{1i} = Rate of change in mathematics (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
β_{00} = Estimated mean intercept, or initial mathematics status of students at year 0
β_{10} = Mean academic year growth rate in mathematics
β_{11} = Mean effect of IB on mathematics 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_{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
r_{1i} = Level 2 residual variance in true rate of change across all individual in the population (between-persons deviation)

4. Results for the 2001-2004 Grade 5-8 Cohort

This chapter describes the results of the descriptive, multivariate, and HLM analyses of the 2001-2004 Grade 5-8 Cohort. Results for the 2002-2004 Grade 8-10 Cohort are provided in Chapter 5.

4.1 Achievement and Demographics

There were 990 students in the 2001-2004 Grade 5-8 Cohort, 12 percent of whom (123) participated at least two years in the middle-school IB program during 2001-2004. For the purposes of these analyses, IB participation is defined as participation for two or more years during middle school. Students who participated only one year during middle school or in the primary program only are not considered as IB participants in middle school. Since information on primary years IB participation also is available for the 1999-2000 and 2000-2001 school years, we are able to track the cohort members' IB participation from grade 4 through grade 8 Table 3 indicates that IB participation ranged from 2 to 5 years during elementary and middle school in 2000 – 2004, with the majority participating 3 years.

Of the 123 students in this cohort, 52 (42%) also participated in the primary years IB program.

Table 3
Number of Years of Participation in District 20's IB Program
2000 – 2004 Grade 5-8 Cohort

Years in IB Program	n	Percent
2	7	5.7
3	66	53.6
4	5	4.1
5	45	36.6
Total	123	100.0

The results of the descriptive analyses and t-tests for differences in means indicate that IB students score significantly higher than non-IB students in mathematics (and in reading) in all four years of this study. However, the rate of growth in mathematics performance for IB students appears to be significantly lower than the growth rate of non-IB students in the district's middle schools by an average of 3.4 scale score points per year. The results of the t-tests are presented in Table 4.

Table 4
Mean Mathematics Achievement and Other Characteristics
of IB and Non-IB Students, Grades 5-8
CSAP 2002 – 2004

Characteristic	IB Students	Non-IB Students	Mean Difference
Mean 5th Grade Mathematics Score, 2001	560	514	46****
Mean 6th Grade Mathematics Score, 2002	602	547	55****
Mean 7th Grade Mathematics Score, 2003	602	568	34****
Mean 8th Grade Mathematics Score, 2004	615	579	36****
Mean Gain in Math Scores, Grade 5-8, 2001-2004	18.4	21.8	- 3.4***
Mean 5th Grade Reading Score, 2001	672	637	35****
Mean 8th Grade Reading Score, 2004	715	677	38****
Percent Female	56	48	8*
Percent Poverty	1	3	- 2**
Percent Minority	18	12	5^{ns}
Percent non-English Language	4	1	3*
No. of Students	123	867	- 744

Significance of Differences: *p = .10 **p = .01 ***p = < .005 **** p = < .001 *ns* = not significant

IB students are slightly more likely than other students to be female, from more advantaged backgrounds, and to be from a non-English language background. It should be noted that the IB program admits any student who requests participation, but in District 20 most of the students in the middle school program come from the Mountain Ridge Middle School catchment area.

4.1.1 Distributional Characteristics

The distributional characteristics of IB/non-IB mathematics achievement indicate that both non-IB and IB students span the range of mathematics scores. Possible mathematics scores on the vertical scale for the grade 5 through 10 CSAP range from 220 to 950. Figures 1a-b and 2a-b illustrate the distributions on the 2001 grade 5 and 2004 grade 8 CSAP mathematics assessments, respectively.

These figures indicate that the grade 5 mathematics scores of non-IB students range from 263 to 797, while the range for IB students is 455 to 728, and that the means are significantly different for the two groups in both years.

Figure 1a
Distribution of Mathematics Scores on the
Grade 5 CSAP, 2001 for
Non-IB Students

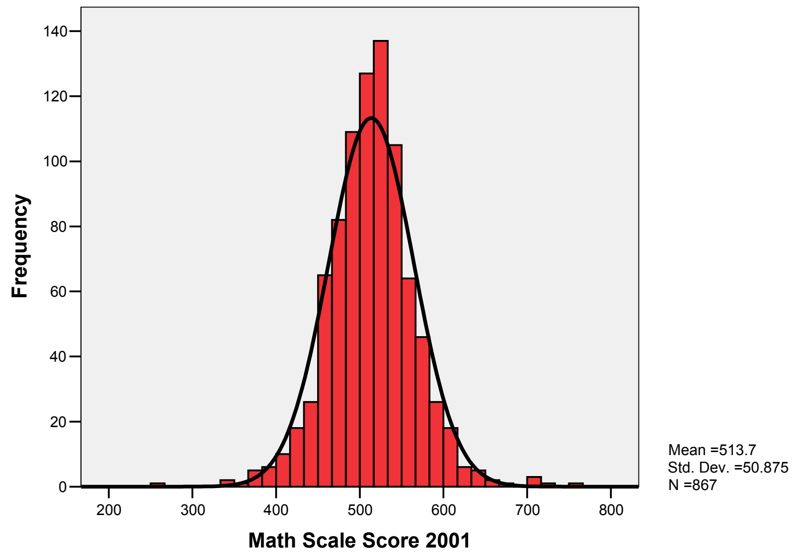
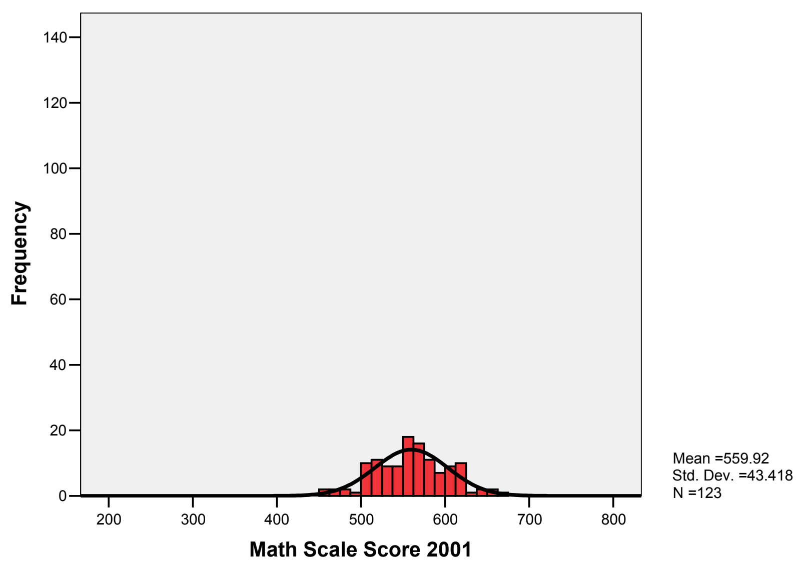


Figure 1b
Distribution of Mathematics Scores on the
Grade 5 CSAP, 2001 for
IB Students



The distributions of scores on the 2004 grade 5 mathematics assessment are similar to those discussed above, with both IB and non-IB students spanning the range of scores with no obvious concentration on scores (e.g., in the upper tail for IB students). The distributions of the grade 8 mathematics scores are illustrated in Figures 2a-2b below.

Figure 2a
Distribution of Mathematics Scores on the
Grade 8 CSAP, 2004 for
Non-IB Students

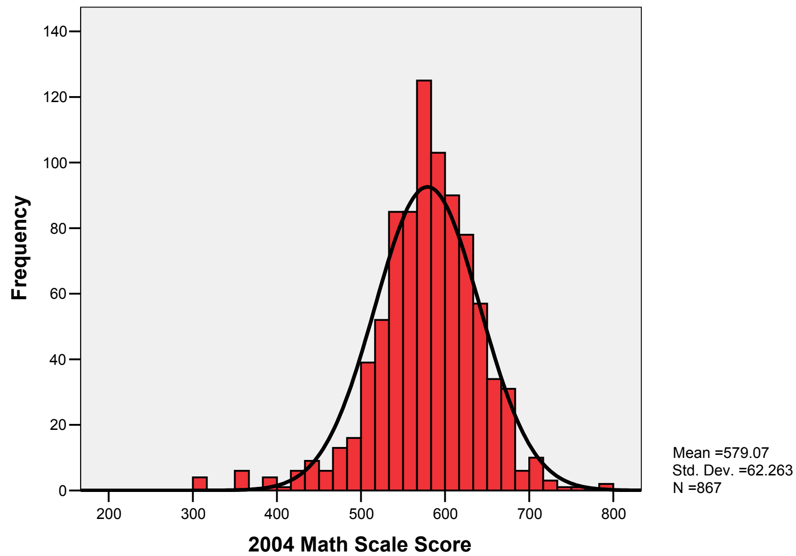
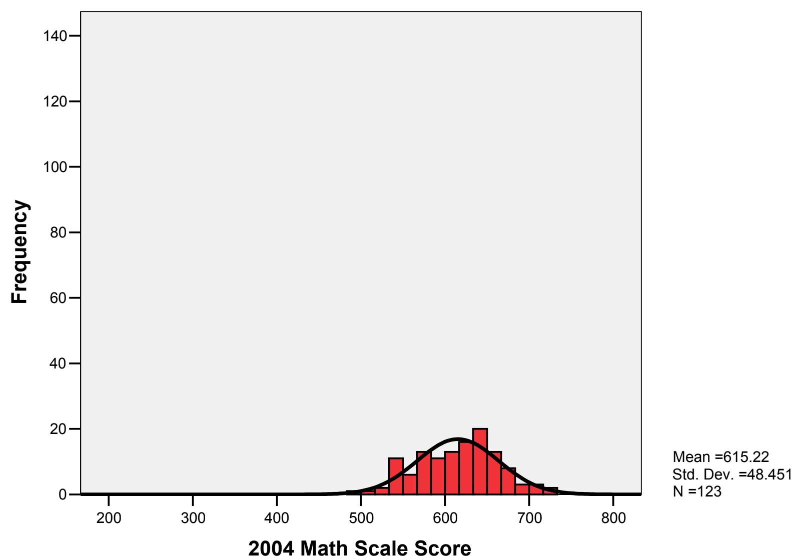


Figure 2b
Distribution of Mathematics Scores on the
Grade 8 CSAP, 2004 for
IB Students



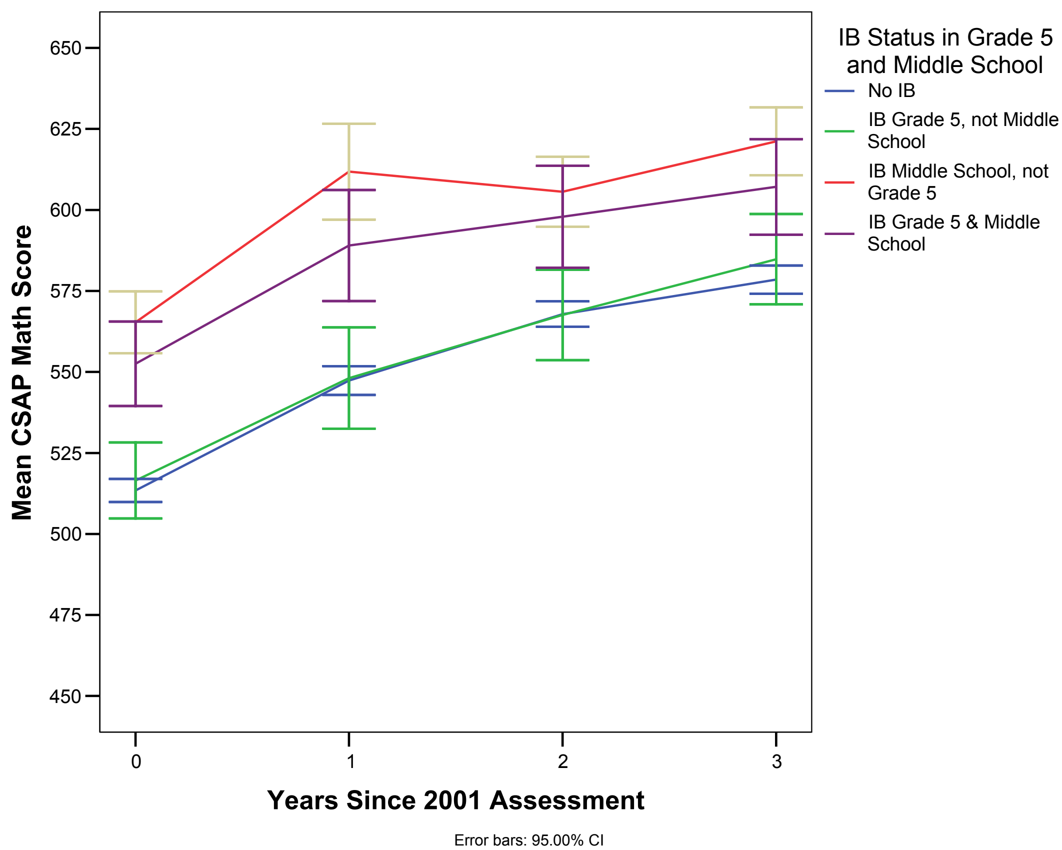
4.1.2 Effects of IB Participation

Regarding effects of IB participation and growth in mathematics achievement over time, it is hypothesized that:

- student participation in the IB program during elementary school is related to performance on the grade 5 math assessment;
- participation during middle school affects math performance and learning rate; and
- students who were in the primary years IB program, but subsequently dropped out of the program in middle school will demonstrate lower math scores than their more IB-persistent counterparts.

The observed, unadjusted data partially confirm the three hypotheses. The observed data indicate that IB students score higher that non-IB students at all four measurements but there appears to be little or no difference in their growth rates. There is no difference in the mathematics test scores of students who participated in IB during middle school only or during both elementary and middle school. Students who dropped out of the IB program by middle school scored no better than students who never participated in the program, as demonstrated in Figure 3 below. The error bars represent the 95 % confidence interval around each mean. Bars that overlap vertically indicate no significant difference between their respective lines.

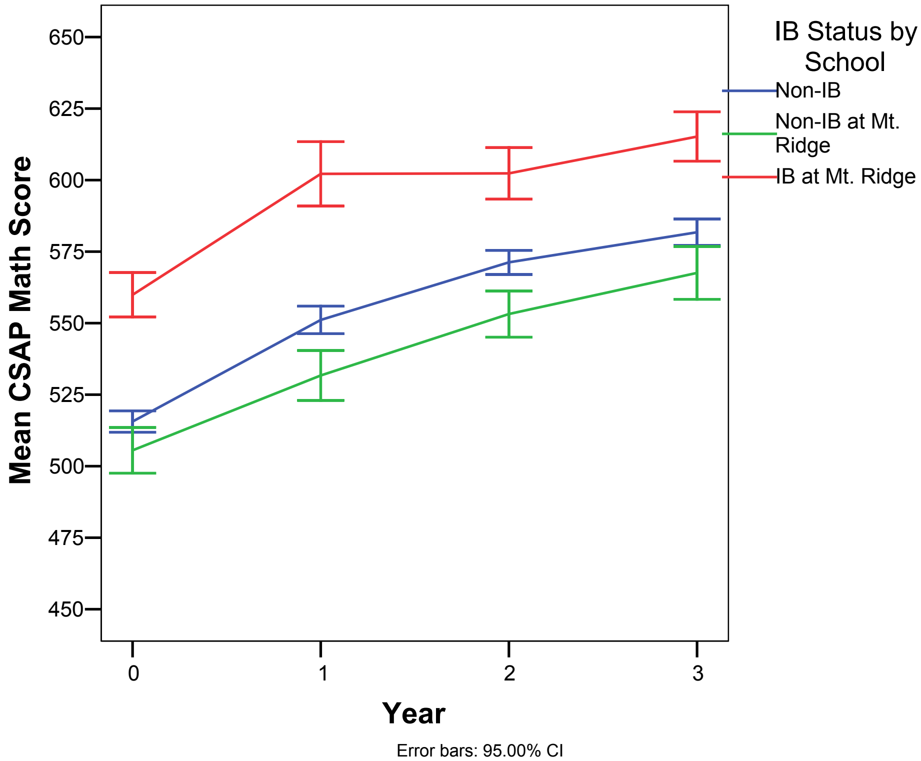
Figure 3
Observed Mean Math Scores by IB Status, Grades 5-8, 2001-2004



The middle years IB program is housed at Mountain Ridge Middle School (MRMS). While the program is open to any student in the district, the majority of IB students are from the Mountain Ridge MS catchment area. The figure below compares the mean CSAP Math Scores for non-IB MRMS students, MRMS IB students, and all other students.

Figure 4

Observed Mean Math Scores by School and IB Status, Grades 5-8, 2001-2004

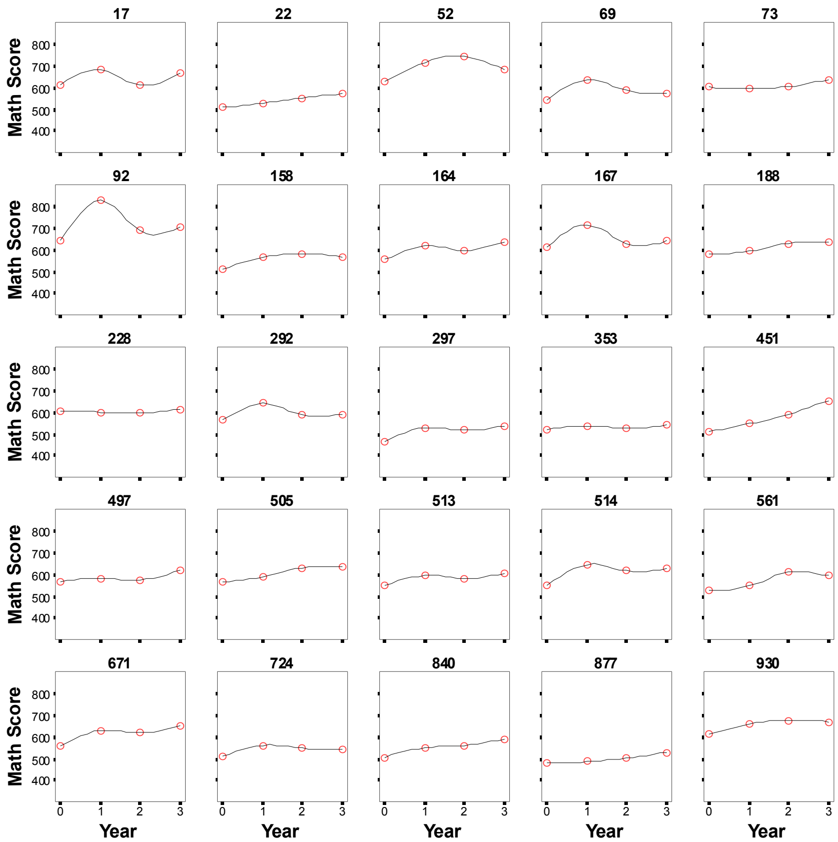


IB students score significantly higher than non-IB students at MRMS and other non-IB students in the district. On the 2004 Grade 8 math assessment (Year 3), Mountain Ridge IB Students out-scored other MRMS students by 47 points and other district students by 33 points.

4.1.3 Empirical Growth Plots

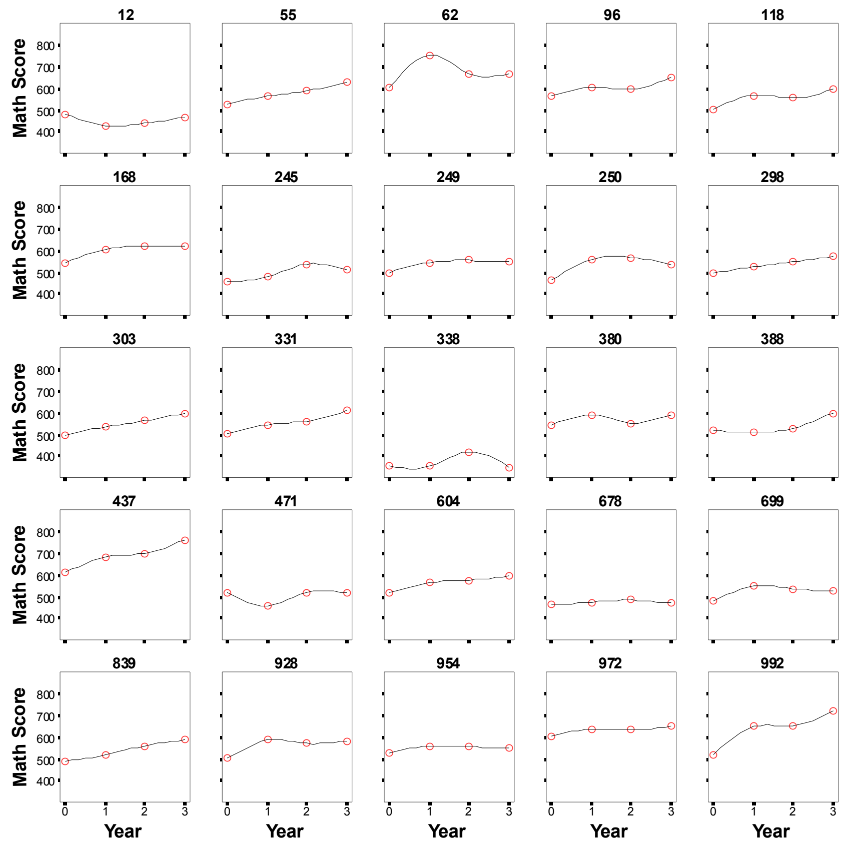
In order to visualize how individuals in this cohort change over time and to determine the best overall parametric model fit (e.g., linear, logistic) to the data, “raw” empirical growth plots are now examined. Random samples of 25 IB and 25 non-IB students are selected and their empirical, nonparametric, growth plots drawn. The “smoothed” empirical growth curves drawn through the actual data are illustrated in Figures 5a and 5b for IB and non-IB students, respectively.

Figure 5a
Nonparametric Mathematics Growth Curves
Grades 5-8, 2001-2004 for a
Sample of IB Students[4]



[4] The "id numbers" above each student plot are assigned at random and cannot be used to identify individual students.

Figure 5b
Nonparametric Mathematics Growth Curves
Grades 5-8, 2001-2004 for a
Sample of Non-IB Students[5]



In these two samples of 25 students, it is apparent that the basic shapes of the curves for both groups are similar. The best common functional form across individual trajectories appears to be linear although there are notable exceptions (e.g., 17, 52 and 92 in the IB sample and 62 and 250 in the non-IB sample).

4.2 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.

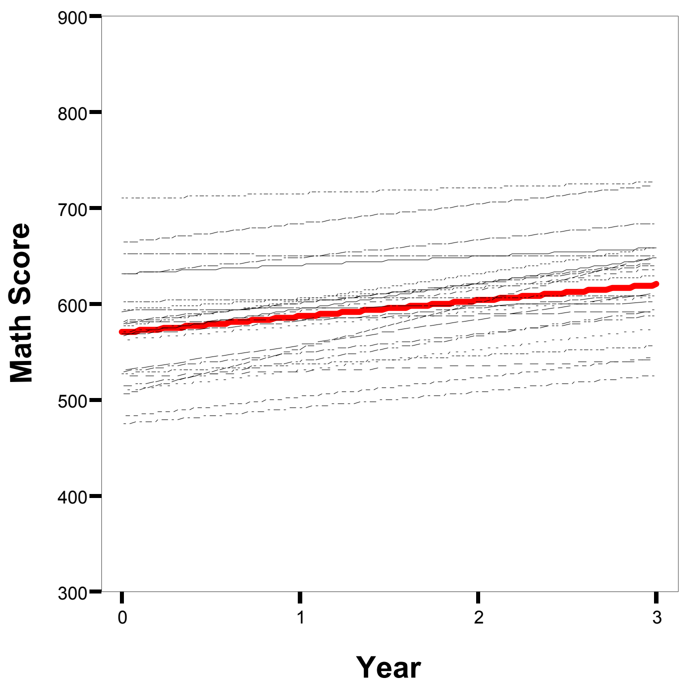
[5] The “id numbers” above each student plot are assigned at random and cannot be used to identify individual students.

4.2.1 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.

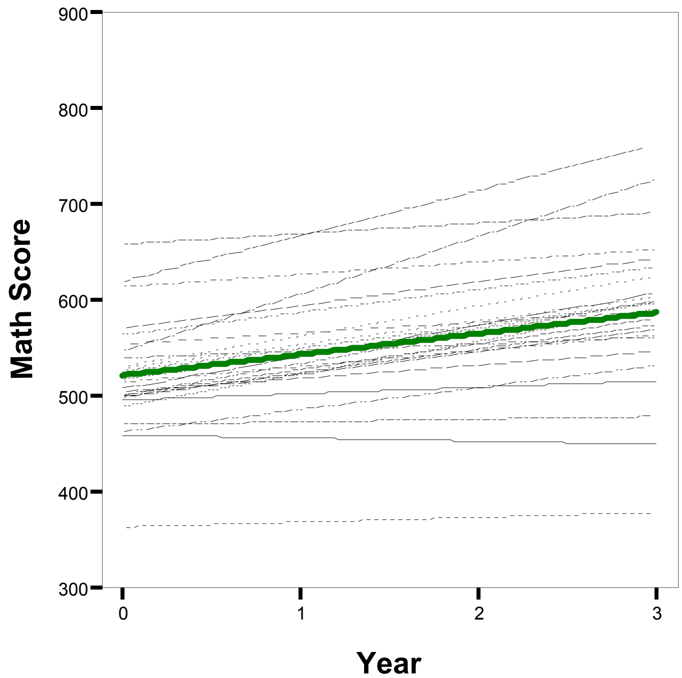
Figures 6a and 6b indicate that there is somewhat more variability in the intercepts (2001 grade 5 test scores) of IB students than for non-IB students. However, the figures also indicate greater variability in the slopes (growth rates) of the non-IB students. The red line in Figure 5a and the green line in Figure 5b represent the average regression lines for the two groups of students. Inspection of the two figures demonstrates the significantly higher mean initial status (mathematics score in Year 0) and the somewhat flatter growth trajectory in the IB sample.

Figure 6a
Ordinary Least Squares Mathematics Growth Trajectories
Grades 5-8, 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 6b
Ordinary Least Squares Mathematics Growth Trajectories
Grades 5-8, 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 7a and 7b 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 mathematics scores.

Figure 7a
Mean OLS Regression Line for Mathematics Growth
Grades 5-8, 2001-2004 for a Sample of IB Students

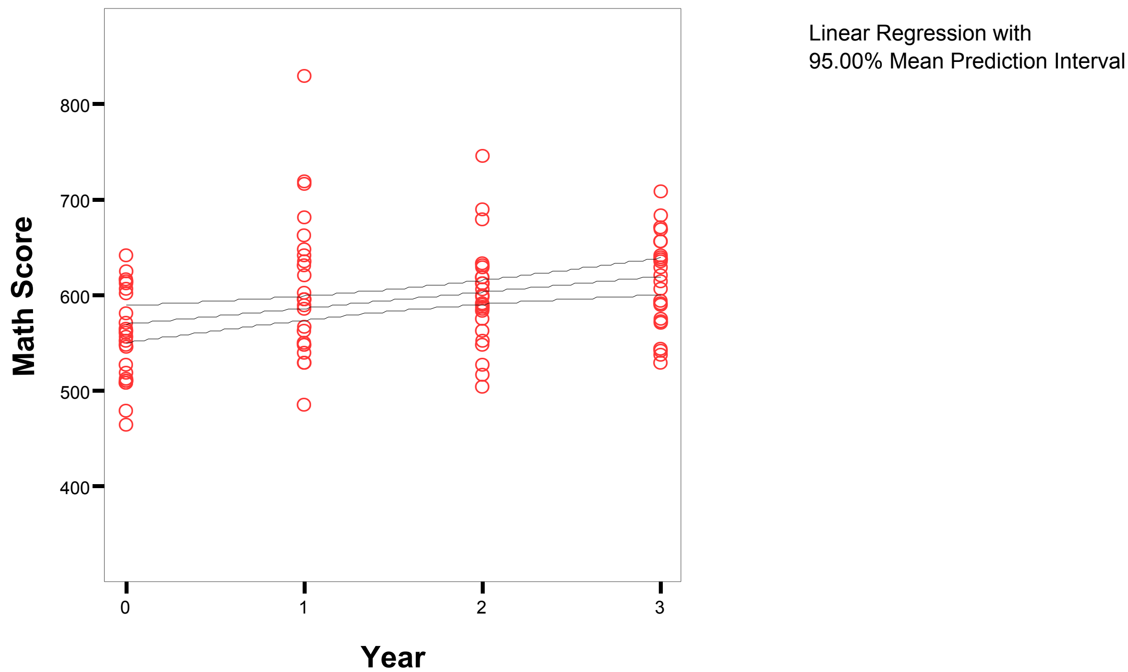
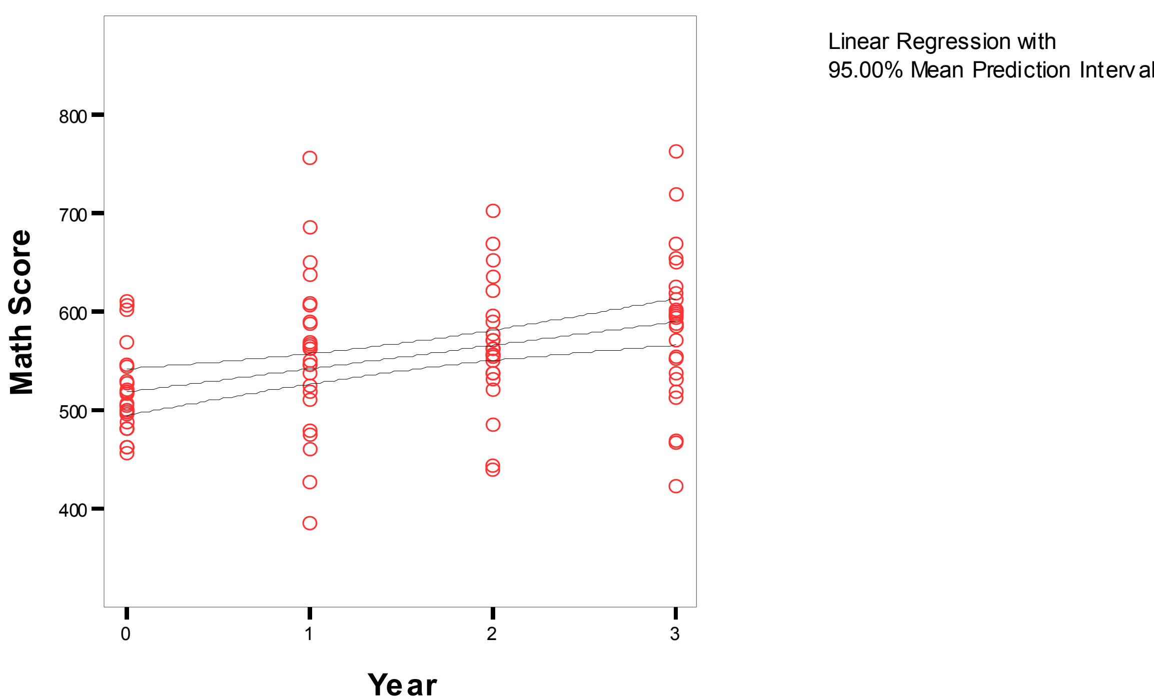


Figure 7b
Mean OLS Regression Line for Mathematics Growth
Grades 5-8, 2001-2004 for a Sample of Non-IB Students



4.2.2 Results from the Multiple Regression: Exploratory Predictive Models

In order to identify factors that are potentially predictive of student achievement and growth in mathematics, bivariate correlations and stepwise linear regression models were estimated prior to the hierarchical linear modeling. The bivariate correlations indicate that the strongest correlates to the outcome, mathematics achievement in eighth grade, are reading scores in grade 5 (r = .67) and in grade 8 (r = .76) and prior mathematics achievement in grade 5 (r = 0.80), grade 6 (r = .87), and grade 7 (r = .89). The strongest correlates to average gain in achievement from fifth through eighth grades are math achievement in grades 6, 7, and 8 (r = .27, .32, and .54, respectively, p = .01). Grade 5 math achievement was only weakly related to average growth in math from grade 5 to grade 7 (r = -.08, p = .05). Better predictors of growth in math achievement appear to be reading performance in grades 5 and 8 (r = .09 and .31, respectively, p = .01).

The preliminary multiple regression analyses indicate no significant effect of IB participation on either mathematics achievement or growth in middle school. The best predictors of math achievement in eighth grade and average gain in achievement are grade 5 math score and grade 8 reading score. The next section describes the results of the hierarchical linear models (HLM) of mathematics achievement and growth during middle school in District 20.

4.3 Results from the HLM Estimation of the Effects of IB Participation on Mathematics Growth in Middle School

The OLS fits to the raw achievement data discussed in § 4.2.1 indicate sufficient variability to warrant more sophisticated modeling. Hierarchical Linear Models (HLM) was used to examine the effects of IB participation on mathematics growth and outcomes. The results of the HLM models are described below.

However, the preliminary regression analyses indicate that none of the available variables, including initial mathematics status and IB participation, are significant predictors of mathematics growth in District 20's middle 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 mathematics assessment.

4.3.1 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, π_{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 mathematics mean plus measurement error.

LEVEL 1 MODEL

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:

$$MATH_{ti} = \beta_{00} + r_{0i} + e_{ti}$$

where

MATH$_{ti}$ is the outcome variable, mathematics score in grade 10

π_{0i} = True math mean for person *i*
β_{00} = Grand mean math 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 5.

4.3.2 Unconditional Growth Model

An unconditional growth model introduces time as the only level 1 predictor and includes no level 2 predictor. 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. By partitioning the variance into within-persons (growth) and between-persons (individual factors, including those relating to selection effects) without controlling for additional person-level variables, the unconditional model provides the baseline for evaluating the adequacy of subsequent models.

Based on the exploratory analyses described in § 4.1.3 and 4.2.1, a linear change trajectory is specified. The Level 1 and Level 2 submodel specifications are shown below:

LEVEL 1 MODEL

$MATH_{ti} = \pi_{0i} + \pi_{1i} (YEAR_{ti}) + e_{ti}$

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:

$MATH_{ti} = \beta_{00} + \beta_{10} (YEAR) + r_{0i} + r_{1i} (YEAR_{ti}) + e_{ti}$

where,

$MATH_{ti}$ is the outcome variable, mathematics score in grade 8

π_{0i} = Initial mathematics status of person *i*, that is, the expected outcome for that student in the spring of grade 5 (when Year = 0)
π_{1i} = Rate of change in mathematics (growth rate) for person *i*
YEAR is 0 at spring 2001, 1 at spring 2002, 2 at Spring 2003, & 3 at Spring 2004 (grades 5, 6, 7, & 8)
β_{00} = Estimated mean intercept, or initial mathematics status, of all students at year 0
β_{10} = Mean academic year growth rate in mathematics
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
r_{1i} = Level 2 residual variance in true rate of change across all individual in the population (between-persons deviation)

The results of the unconditional means and growth models are provided in Table 5.

Table 5
Comparison of Unconditional Means and Growth Models
of Mathematics Performance and Growth
Grades 5-8, 2001-2004

		Parameter	Unconditional Means Model	Unconditional Growth Model
Fixed Effects				
Initial status, π_{0i}				
Mean grade 5 math score		β_{00}	557.16*** (1.78)	525.73*** (1.75)
Rate of change, π_{1i}				
Mean student growth rate		β_{10}		20.96*** (0. 38)
Variance Components				
Level 1	Within-persons	e_{ti}	1351.91	575.23
Level 2	For initial status	r_{0i}	2784.44***	2619.49***
	For rate of change	r_{1i}		25.03***
	Correlation	r_{01}		0.39
Goodness-of-fit				
Deviance			41858.47	39481.79
χ^2 statistic				2376.67***

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, π_{0i} (person i's average mathematics score regardless of time). Thus, the intercept represents the average of the grade 5 – 8 math scores averaged across all students. 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 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 (ρ) obtained from the unconditional means model is:

$$\rho = r_{0i} / (r_{0i} + e_{ti}i)$$
$$\rho = 2784.44 / (2784.44 + 1351.91) = 0.3268$$

Thus, approximately one-third (33 %) of the total variation in mathematics performance is attributable to differences between students. These results indicate that additional modeling may be warranted.

When a temporal component is included in the model (i.e., the slope of the growth trajectory), the intercept represents mean student scores at the initial measurement (i.e., grade 5). Thus, students average about 526 on the initial grade 5 mathematics assessment (β_{00}) and grow at an average rate of about 21 scale score points per year (β_{10}) from fifth through eighth grades. The addition of the temporal variable (unconditional growth model) significantly reduces the amount of residual (i.e., unexplained) variance at Level 1 (the growth trajectory) by 57 percent, signifying that there is true growth in mathematics between fifth and eighth grades:

$$[(1351.91 - 575.23) / 1351.91] = .575.$$

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 mathematics status) and rate of change (mathematics growth). Several conditional models are estimated. The initial conditional models indicate that none of the hypothesized covariates (e.g., initial status, gender) significantly predict mathematics growth once IB variables are accounted for. The conditional model described below is the final model and evaluates the effect of middle school IB participation (for 2 or more years) on mathematics achievement and growth from fifth through eighth grades. This model conceptualizes IB as a "treatment" effect.

4.3.3 Final Conditional Model

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 sufficiently "explain" all the observed variation in mathematics growth. Many models were estimated, but goodness-of-fit statistics indicate that models including IB participation are no better than the unconditional growth model.[6]

Several exploratory models estimated in this iterative process indicate that growth in math from grade 5 to 8 does not appear to be related to:

- Initial math status (achievement in Grade 5);
- Reading performance in grade 5; or
- IB participation in grade 5.

[6] However, that apparent lack of improvement in model fit may be an artifact of a change in HLM's default estimation procedure from version 5.0 to 6.0. This possibility warrants further investigation. The models described in § 4.3.3 will be re-estimated prior to preparation of a manuscript for journal publication.

The results for the best fitting conditional model are provided in Table 6. This model examines the effects of grade 5 reading performance on grade 5 mathematics achievement and effects of IB participation in middle school on mathematics growth rate (slope) during middle school without controlling for other student characteristics. The Level 1 and 2 submodels are:

LEVEL 1 MODEL

$MATHEMATICS_{ti} = \pi_{0i} + \pi_{1i} (YEAR_{ti}) + e_{ti}$

LEVEL 2 MODEL

$\pi_{0i} = \beta_{00} + \beta_{01} (READ01) + r_{0i}$

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

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

$MATHEMATICS_{ti} = \beta_{00} + \beta_{01} (READ01) + \beta_{10} (YEAR) + \beta_{11} (IB*YEAR)) + r_{0i} + r_{1i} (YEAR_{ti}) + e_{ti}$

where,

$MATHEMATICS_{ti}$ is the outcome variable, mathematics score in grade 8

π_{0i} = Initial mathematics status of person *i*, that is, the expected outcome for that student in the spring of grade 5 (when Year = 0)
π_{1i} = Rate of change in mathematics (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 5, 6, 7, & 8)
β_{00} = Estimated mean intercept, or initial mathematics status of students at year 0
β_{01} = Mean effect of reading performance on the intercept (initial mathematics status)
β_{10} = Mean academic year growth rate in mathematics
β_{11} = Mean effect of IB on mathematics 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 in middle school
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 individuals in the population
r_{1i} = Level 2 residual variance in true rate of change across all individuals in the population (between-persons deviation)

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

This model attempts to control for possible selection effects by controlling for the effect of reading ability on initial math status (grade 5 math score)[7]. Reading performance, as well as prior performance in other subjects, may be interpreted as incorporating the effects of the numerous personal and contextual factors that influence academic achievement and gains in achievement, including prior learning, opportunities to learn, extrinsic and intrinsic motivation, parental desires, home and community environment, etc. By attempting to control for reading achievement, we are making an effort to include the effects of other unmeasurable and unknowable factors.

Table 6
Effects of Reading Performance and IB Participation on
Mathematics Performance and Growth, Grades 5-8, 2001-2004

		Parameter	Coefficient
Fixed Effects			
Initial status, π_{0i}			
Mean grade 5 math score		β_{00}	525.73*** (1.15)
Mean reading effect		β_{01}	0.83*** (0.02)
Rate of change, π_{1i}			
Mean growth rate for non-IB students		β_{10}	21.44*** (0. 861)
IB effect on growth rate (IB – non-IB gap)		β_{11}	-3.83** (1.03)
Variance Components			
Level 1	Within-persons	e_{ti}	573.39
Level 2	For initial status	r_{0i}	914.87***
	For rate of change	r_{1i}	23.79***
Goodness-of-fit			
Deviance			38627.83
χ^2 statistic			853.96 *ns*

Numbers in parentheses represent standard errors.
*** $p < .001$ ** $p = .001$ *ns* = not statistically significant

The variance decomposition provided in the bottom panel of Table 6 indicates that students' reading level in grade 5 accounts for 65 percent of the residual variance in grade 5 mathematics

[7] Another model attempted to evaluate selection effects on mathematics growth by estimating the effects of reading level on math growth rates, but the effect was close to zero and not significant.

score.[8] In other words, about 65 percent of the variability in students' grade 5 math performance can be "explained" by their reading ability at the time of the assessment. However, it appears that the addition of IB participation explains only about five percent of the residual between-persons variance in rate of change.[9]

On average, students in this cohort scored about 526 points (β_{00}) on the 2001 grade 5 CSAP mathematics assessment.

There are two very striking results shown in Table 6:

1. **On average, students in the IB program demonstrate lower growth in mathematics (β_{11}) in middle school than non-IB students. On average, non-IB students gain about 21 points (β_{10}) annually from grade 5 through grade 8, but IB student average about 4 points less, gaining just over 17 points per year.**
2. **Students' mathematics performance in grade 5 is strongly related to their reading ability (as assessed on the grade 5 CSAP). For every extra score point on the reading assessment, students, on average, score about eight-tenths (0.8) of a point higher on the grade 5 math assessment (β_{01}).**

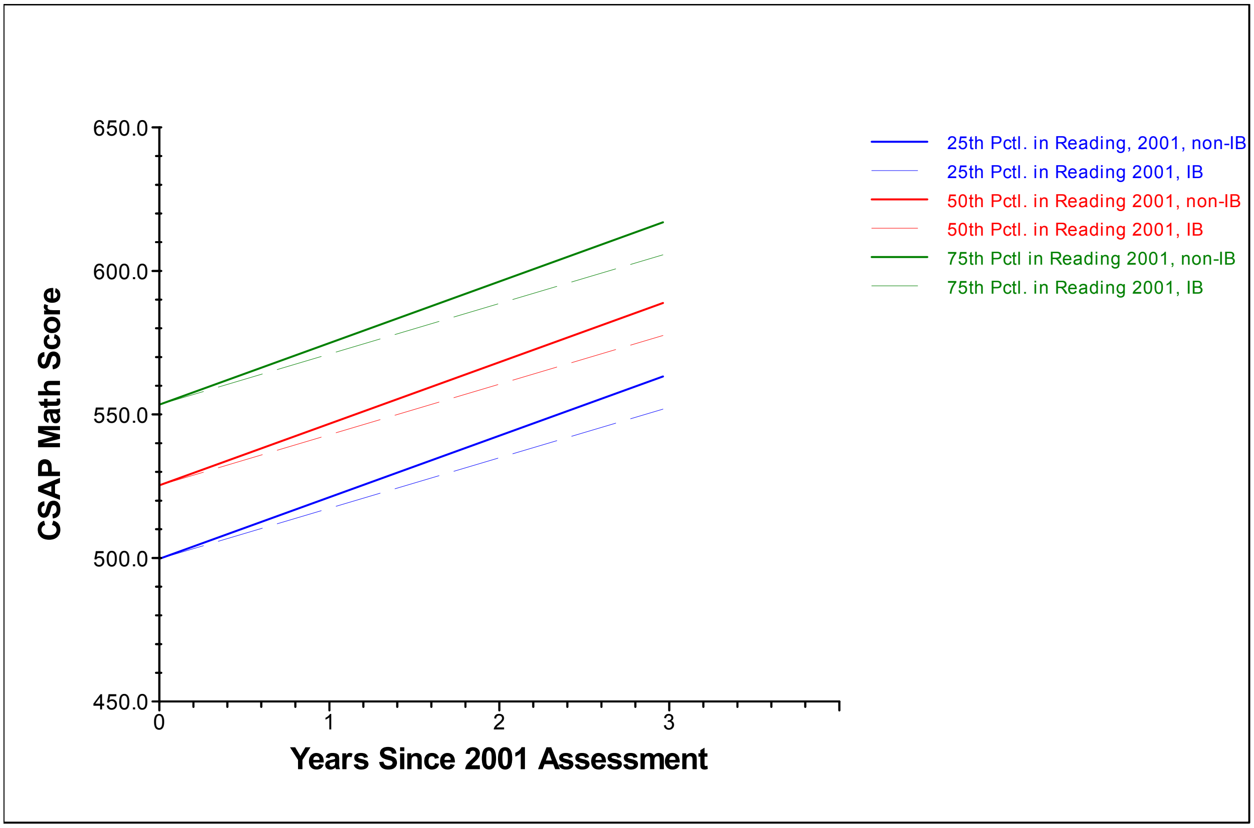
Thus, it appears that IB students generally score higher than non-IB students; however, the differences may be due to selection effects.

For illustrative purposes, the HLM growth trajectories of three groups of similarly-scoring IB and non-IB students are compared in Figure 8. The three groups are: students who score at the 25th percentile on the grade 5 reading assessment, students who score at the 50th percentile (median), and students who score at the 75th percentile on the grade 5 reading assessment.

Note that the growth rates of IB students, denoted by the dashed lines, are significantly flatter than those of the non-IB students at the three percentile groupings. Prior research in this district (as well as in other Colorado districts and other states) indicates that growth rates typically are *inversely* related to initial status. However, in this cohort of students, growth rate and initial status are *positively* related: r = 0.39. The association is fairly weak, as demonstrated by the small value of r, but the positive direction suggests that the higher a student's grade 5 mathematics score, the higher the growth rate.

[8] The percent reduction in residual variance for the intercept that is between persons is estimated as percentage change from the unconditional model: [(2619.49 – 914.87) / 2619.49] = 0.651.
[9] I.e., [(25.03 – 23.79) / 25.03] = 0.0495.

Figure 8
Modeled Mathematics Growth Trajectories of IB and Non-IB Students who Score at the 25th, 50th, & 75th Percentiles in Grade 5 Reading
Grades 5-8, 2001-2004



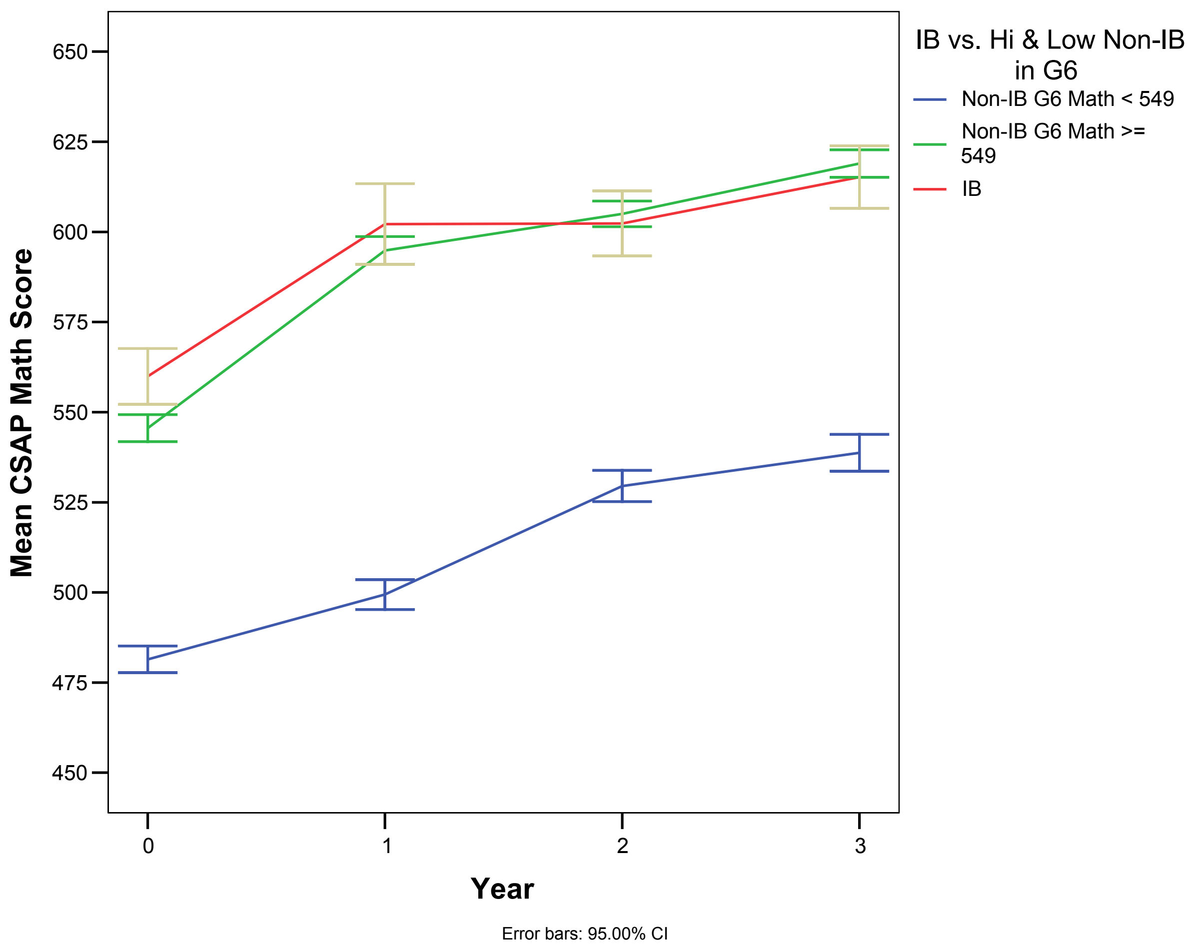
4.3.4 Comparison of IB and High-Scoring Non-IB Students

In further attempts to mitigate the effects of selection bias (if any) and control for mathematics achievement, IB students are compared with non-IB students who score at or above their median sixth grade math score of 549. The grade six score, rather than the grade 5 score, was selected because we want to have as comparable a group as possible. The grade 5 assessment is the first CSAP math test students have taken, and students typically score lower than their actual abilities would indicate on the first experience with a particular type of standardized test. Further, both the IB and non-IB students have completed one year of middle school and have matured somewhat. The median grade 6 math score is used simply as a cut to separate lower-performing from higher performing non-IB students, and is not used as a predictor in the analyses described below.

Figure 9 illustrates the observed performance on the grades 5 – 8 math assessments of IB students, non-IB students who score below 549 on the CSAP, and non-IB students who score at or above the median of 549. The results for the lower performing non-IB students are shown for information purposes only. This group of students is excluded from the remainder of the analyses comparing the achievement of IB and other higher performing students.

Note that the growth rate of the higher performing non-IB students appears to be slightly steeper than that of the IB students. Their grade 5 initial status is significantly lower than that of the IB students, but by grades 7 and 8 their test scores equal those of the IB students. The error bars represent the 95 % confidence interval around each mean. Bars that overlap vertically indicate no significant difference between their respective lines.

Figure 9
Mean Math Scores for IB Students and for Non-IB Students Scoring Above or Below 549 (median) on the Grade 6 CSAP Math Assessment
IB vs. High & Low Non-IB in Grade 6



In order to gain further insight into the effect of IB participation over time within a high-achieving cohort, IB students and only the higher scoring non-IB students are compared via a series of HLM models. If IB participation has a positive effect on the learning rates of students, we would expect to see a steeper slope for IB students among higher achieving students in general. The final model estimated is the same as for the total group, whose results are shown in Table 5 and

Figure 7. To reiterate, the model attempts to control for possible selection effects by controlling for the effect of reading ability on initial math status (grade 5 math score). The Level 1 and 2 submodels are:

LEVEL 1 MODEL

$$MATHEMATICS_{ti} = \pi_{0i} + \pi_{1i} (YEAR_{ti}) + e_{ti}$$

LEVEL 2 MODEL

$$\pi_{0i} = \beta_{00} + \beta_{01} (READ01) + r_{0i}$$

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

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

$$MATHEMATICS_{ti} = \beta_{00} + \beta_{01} (READ01) + \beta_{10} (YEAR) + \beta_{11} (IB*YEAR)) + r_{0i} + r_{1i} (YEAR_{ti}) + e_{ti},$$

where the model parameters and residuals are described as above.

In this model, the intercept (β_{00}) is the grand mean for initial status (grade 5 mathematics mean score) across all students in this analysis, while the parameter, β_{01}, represents the effect of reading performance on math score. The β_{11} coefficient on the slope (rate of change) represents the effect of IB participation on mathematics growth rate (β_{10}). Students in this high achieving half of the cohort scored, on average, about 560 points on the 2001 grade 5 CSAP mathematics assessment (β_{00}) and averaged a growth rate of about 23 scale score points per year (β_{10}) from fifth through eighth grades. The results are presented in Table 7 and Figure 9.

The most striking result is that within this high achieving group, the negative IB – non-IB gap is even greater, - 5.21, than the gap when the entire cohort is considered (- 3.83). In other words, among higher achieving students in District 20's middle schools, IB students gain, on average, only about 17 scale score points per year, as opposed to nearly 23 points gained by higher achieving non-IB students.

Table 7
Effects of Reading Performance and IB Participation on Mathematics Performance and Growth of Higher Performing Students, Grades 5-8, 2001-2004

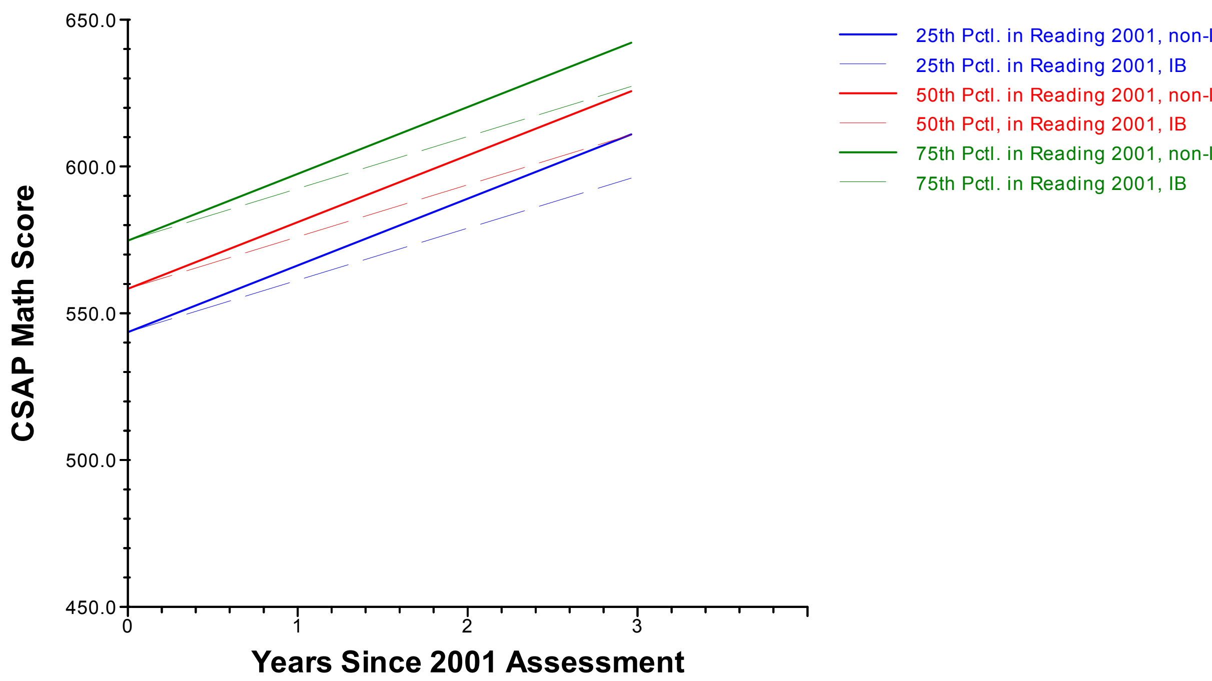
		Parameter	Coefficient
Fixed Effects			
Initial status, π_{0i}			
Mean grade 5 math score		β_{00}	559.52***
			(1.36)
Mean reading effect		β_{01}	0.57***
			(0.03)
Rate of change, π_{1i}			
Mean growth rate for non-IB Students		β_{10}	22.73***
			(0. 52)
IB growth rate effect (IB – non-IB gap)		β_{11}	- 5.02***
			(1.07)
Variance Components			
Level 1	Within-persons	e_{ti}	645.24
Level 2	For initial status	r_{0i}	757.46***
Goodness-of-fit			
Deviance			21761.50
χ^2 statistic			271.57***

Numbers in parentheses represent standard errors. *** $p < .001$

These results are illustrated in Figure 9.

As in Figure 8, the HLM mathematics growth trajectories of three groups of similarly-scoring IB and higher-scoring non-IB students are compared in Figure 10. The three groups are: students who score at the 25th percentile on the grade 5 reading assessment, students who score at the 50th percentile (median), and students who score at the 75th percentile on the grade 5 reading assessment.

Figure 10
Modeled Mathematics Growth Trajectories of IB and High-Scoring Non-IB Students who Score at the 25th, 50th, & 75th Percentiles in Grade 5 Reading
Grades 5-8, 2001-2004



Note that even in this high-achieving portion of the cohort, the growth rates of IB students, denoted by the dashed lines, are significantly flatter than those of the non-IB students at the three percentile groupings. The value in considering the information illustrated by Figure 9 is that controlling for starting point, IB students appear to progress from grade 5 through 8 at slightly *slower* rates than non-IB students.

4.4 Summary of Results for the 2001-2004 Grade 5-8 Cohort

In summary, the analyses reported above indicate IB students score significantly higher on the grade 5 mathematics assessment than non-IB students. However, the growth rates of IB students during middle school are significantly lower than those of other students in general or higher-scoring non-IB students.

Conservatively speaking, the IB program (as displayed by this study) does not appear to provide an obvious acceleration to students' learning growth in mathematics during middle school. However, we do not have sufficient data to adequately explain selection effects (e.g., student motivation; parental encouragement to enroll in IB, which might result in average or bright (but not the brightest) students enrolling in IB). We also cannot know how these IB students would have performed without the IB program. Such an evaluation can only be made through a randomized experimental design, in which students are randomly assigned to the IB or regular education program, irrespective of parental desires, students' prior achievement, students' desires, student demographics, geography, etc.

5. Results for the 2002-2004 Grade 8-10 Cohort

This chapter describes the results of the descriptive, multivariate, and HLM analyses of the 2002-2004 Grade 8-10 analytical cohort

5.1 Achievement and Demographics

There were 1133 students in the 2002-2004 Grade 8-10 Cohort, 6.4 percent of whom (72) participated in the IB program in tenth grade in the 2003-2004 school year. Students who participated in the program in ninth grade but not in tenth are classified as "non-IB." Since information on IB participation is available for the 2000-2001 school year, we are able to track the cohort members' IB participation from grade 7 through grade 10. Table 8 indicates that IB participation by this analytical cohort ranged from 2 to 4 years during the period 2000 – 2004, with the vast majority (80 %) participating four years.

Table 8
Number of Years of Participation in District 20's IB Program
2001 – 2004 Grade 8-10 Cohort

Years in IB Program	n	Percent
2	5	6.9
3	9	12.5
4	58	80.6
Total	72	100.0

An additional forty-three students participated in IB at least one year during the period 2000-2003, but did not participate in 2003-2004 school year.

The results of the descriptive analyses and t-tests for differences in means indicate that IB students in the diploma program score significantly higher than non-IB students in mathematics (and in reading) in all three years of this study. In addition, the rate of growth in mathematics performance for IB students appears to be significantly higher than the growth rate of non-IB students in the cohort by an average of 5.4 scale score points per year. The results of the t-tests are presented in Table 9.

Table 9
Mean Mathematics Achievement and Other Characteristics
of IB and Non-IB Diploma Students, Grade 8-10
CSAP 2002 – 2004

Characteristic	IB Students	Non-IB Students	Mean Difference
Mean 8th Grade Mathematics Score, 2002	630	582	48**
Mean 9th Grade Mathematics Score, 2003	651	597	54**
Mean 10th Grade Mathematics Score, 2004	664	605	53**
Mean Gain in Math Scores, Grade 5-8, 2001-2004	17	12	5**
Mean 8th Grade Reading Score, 2002	722	676	46**
Mean 10th Grade Reading Score, 2004	753	703	50**
Percent Female	56	48	8
Percent Poverty	0	2	- 2**
Percent Minority	24	13	11 *
Percent non-English Language	1	2	- 1 *ns*
No. of Students	72	1061	- 989

Significance of Differences: *p = .05 ** p = < .001 *ns* = not significant

IB students are more likely than other students to be female, from more advantaged backgrounds, and an ethnic minority. It should be noted that the IB program admits any student who requests participation, but in District 20 most of the students in the high school program come from the Rampart High School catchment area.

5.1.1 Distributional Characteristics

The distributional characteristics of IB/non-IB mathematics achievement indicate that both non-IB and IB students span the range of mathematics scores. Possible mathematics scores on the vertical scale for the grade 5 through 10 CSAP range from 220 to 950. Figures 11a-b and 12a-b illustrate the distributions on the 2002 grade 8 and 2004 grade 10 CSAP mathematics assessments, respectively.

Figures 11a and 11b indicate that the grade 8 mathematics scores of non-IB students range from 310 to 890, while the range for IB students is 500 to 717, and that the means are significantly different for the two groups.

Figure 11a
Distribution of Mathematics Scores on the
Grade 8 CSAP, 2002 for
Non-IB Students

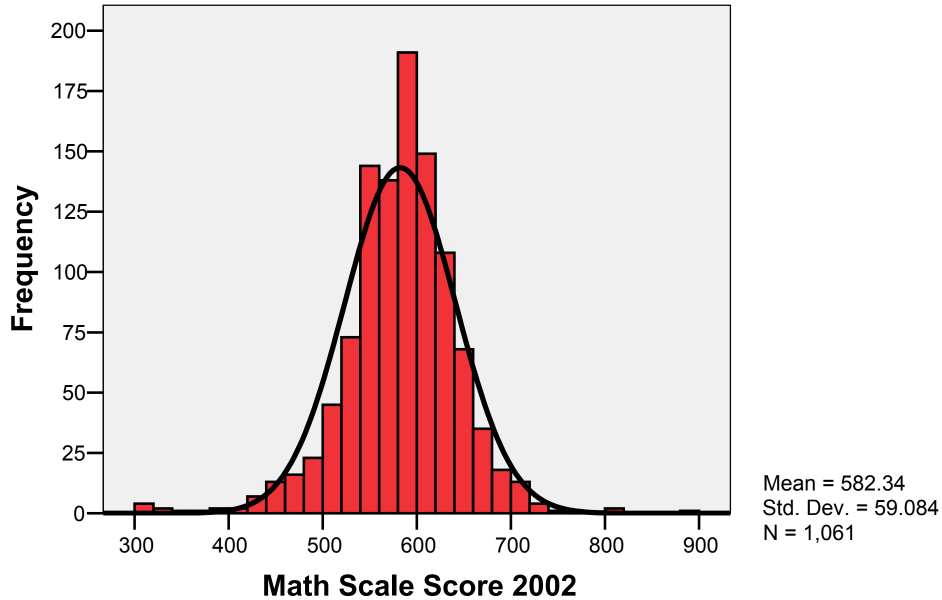
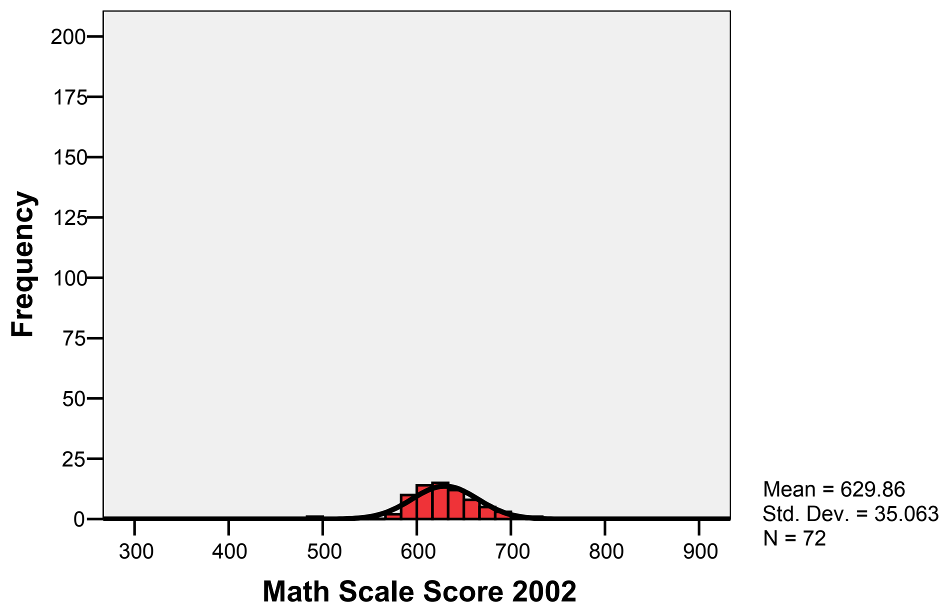


Figure 11b
Distribution of Mathematics Scores on the
Grade 8 CSAP, 2002 for
IB Students



The distributions of scores on the 2004 grade 10 math assessment show that non-IB students demonstrate a lower average mean score and a greater concentration of scores in the lower tail than do IB students.

Figure 12a
Distribution of Mathematics Scores on the
Grade 10 CSAP, 2004 for
Non-IB Students

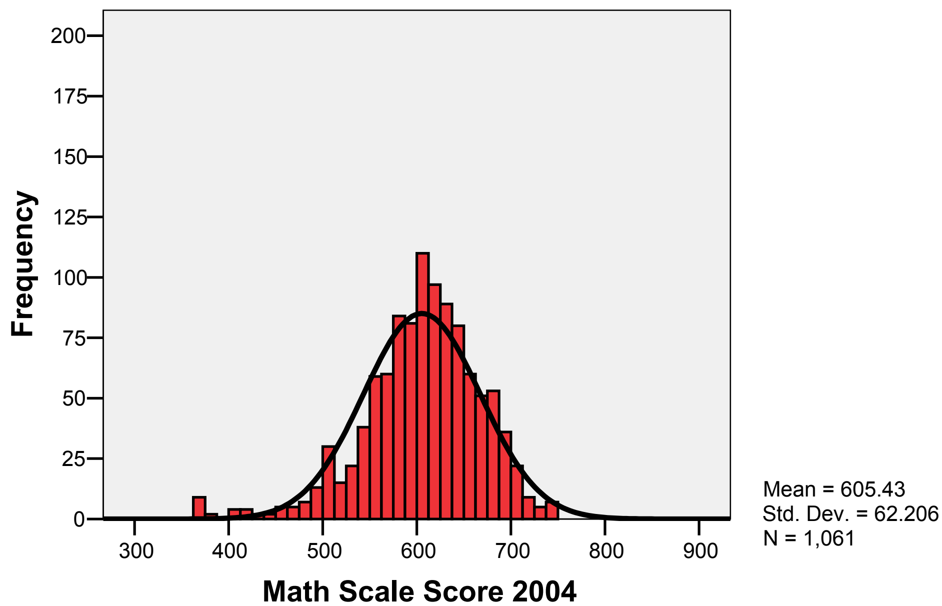
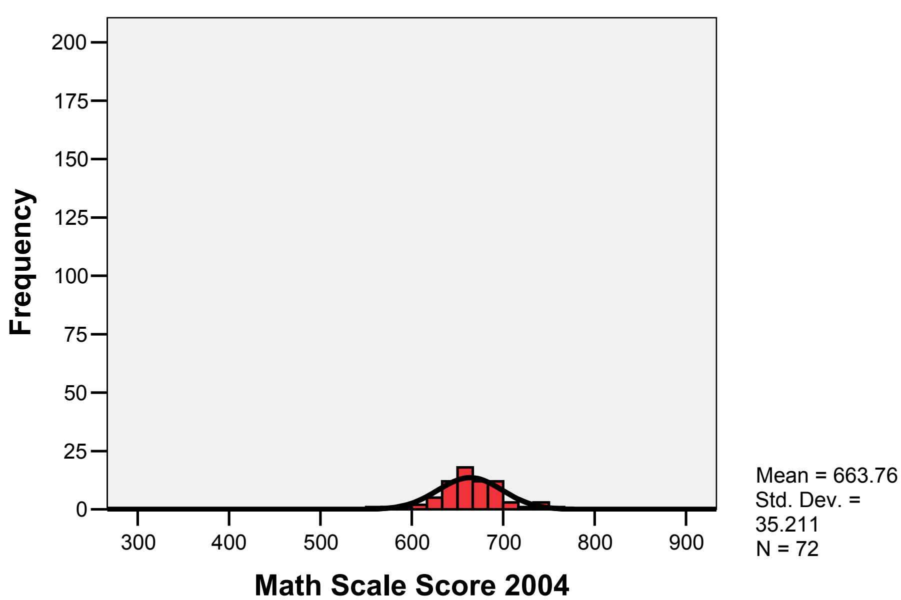


Figure 12b
Distribution of Mathematics Scores on the
Grade 10 CSAP, 2004 for
IB Students



5.1.2 Effects of IB Participation

Regarding effects of IB participation and mathematics achievement, it is hypothesized that:

- student participation in the IB program during high school is related to performance on the grade 10 math assessment;
- participation during high school affects math learning rate; and
- students who were in the middle years IB program, but left the program in high school, will demonstrate lower math scores than their more persistent counterparts.

The observed, unadjusted data confirm the three hypotheses. The observed data indicate that IB students score higher that non-IB students at all three measurements. There is no difference in the mathematics test scores of students who participated in IB during diploma only or both middle and diploma programs. Students who participated in the middle years program, but dropped out in high school, still score better than students who never participated in the program, as demonstrated in Figure 13 below. In addition, students in the diploma program show greater gains than other students. The error bars represent the 90% confidence interval around each mean. Bars that overlap vertically indicate no significant difference between their respective lines.

Figure 13
Observed Mean Math Scores by IB Status, Grades 8-10, 2002-2004

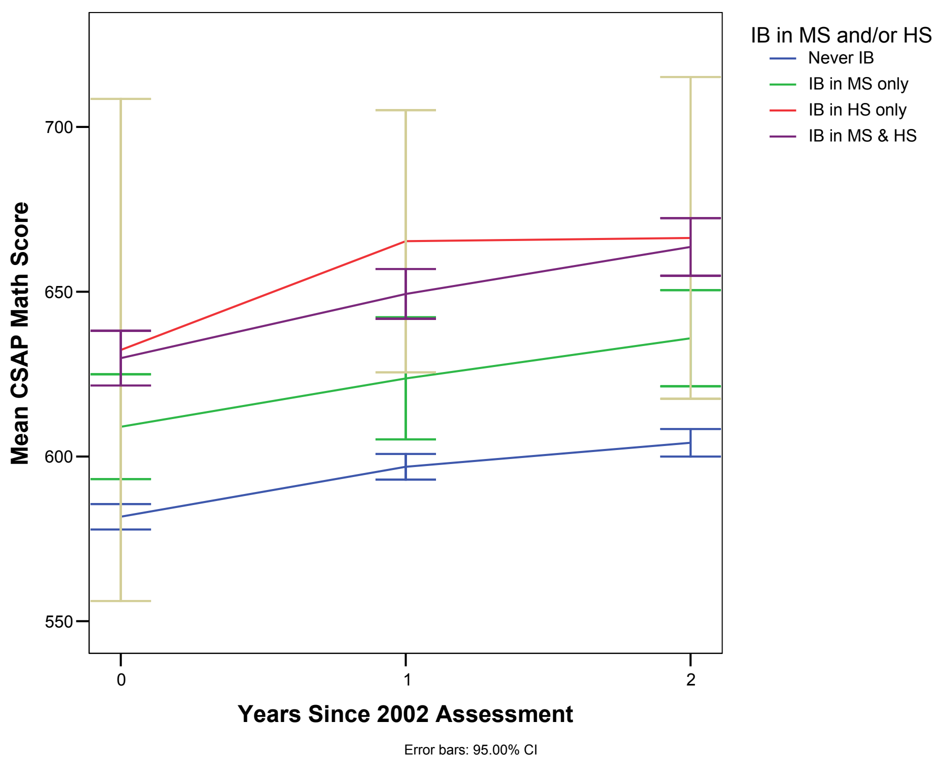
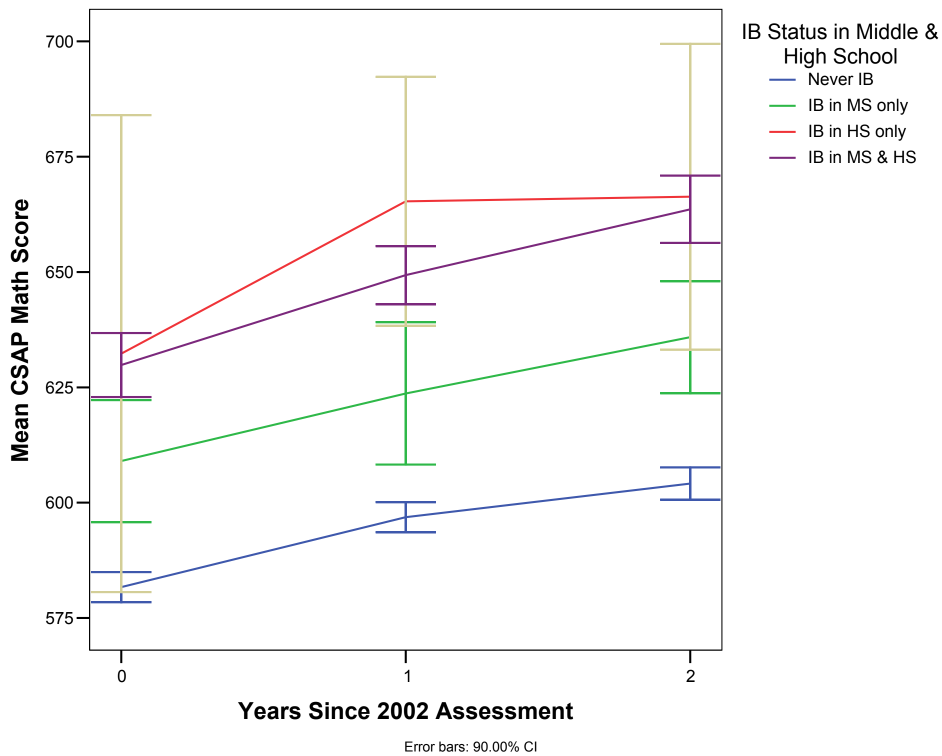


Figure 14 compares the 2002 through 2004 mathematics performance of students who are in the diploma IB program (Rampart High School only), other students at Rampart (RHS) and the remainder of students in the district.

Figure 14
Observed Mean Math Scores by IB Status and School, Grades 8-10, 2002-2004



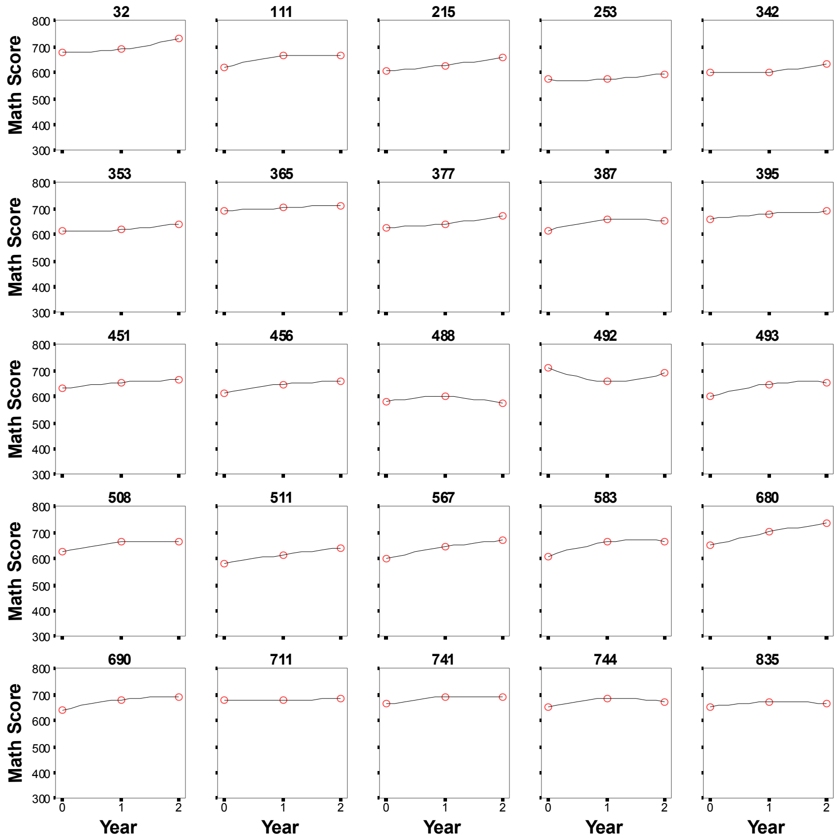
The IB students significantly outscore non-IB students, and there is very little difference between non-IB students at RHS and other non-IB students in the district.

5.1.3 Empirical Growth Plots

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

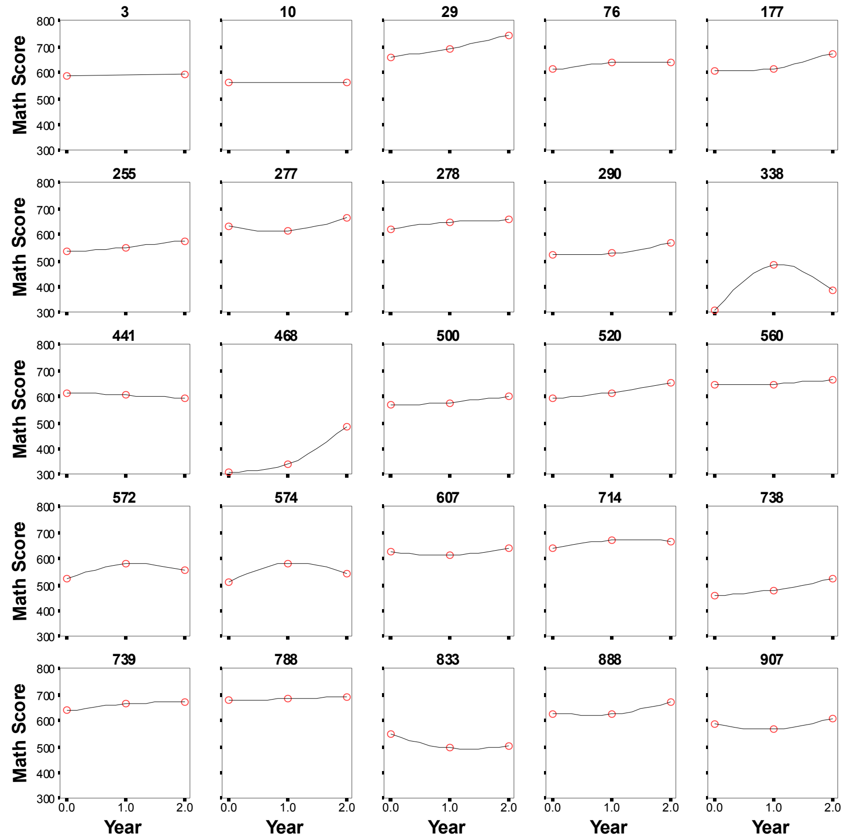
The "smoothed" empirical growth curves drawn through the actual data are illustrated in Figures 15a and 15b for IB and non-IB students, respectively.

Figure 15a
Nonparametric Mathematics Growth Curves
Grades 8-10, 2002-2004 for a
Sample of IB Students[10]



10 The "id numbers" above each student plot are assigned at random and cannot be used to identify individual students.

Figure 15b
Nonparametric Mathematics Growth Curves
Grades 8-10, 2002-2004 for a
Sample of Non-IB Students[11]



In these two samples of 25 students, it is apparent that the basic shapes of the curves for both groups are similar, although the IB curves appear to be slightly more linear. Thus, the best common functional form across individual trajectories appears to be linear although there are notable exceptions (e.g., 492 in the IB sample and 338 and 574 in the non-IB sample).

5.2 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.

---

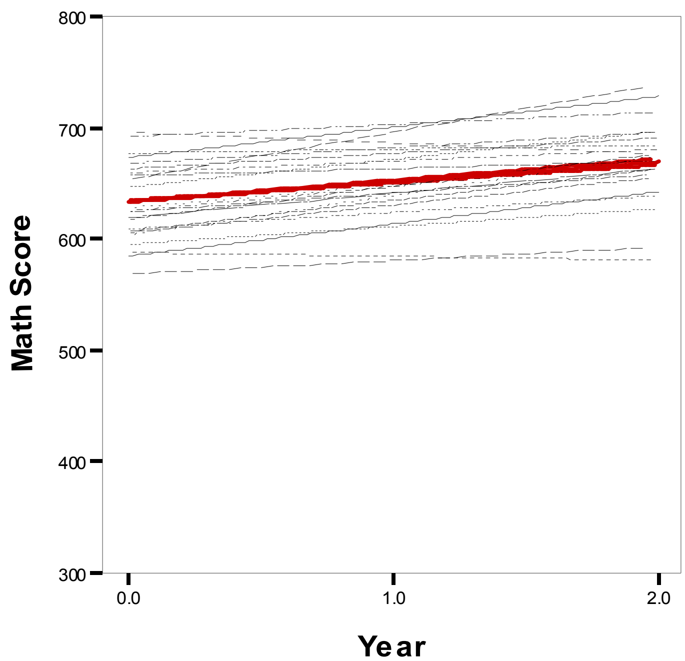
[11] The "id numbers" above each student plot are assigned at random and cannot be used to identify individual students.

5.2.1 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 16a and 16b, respectively.

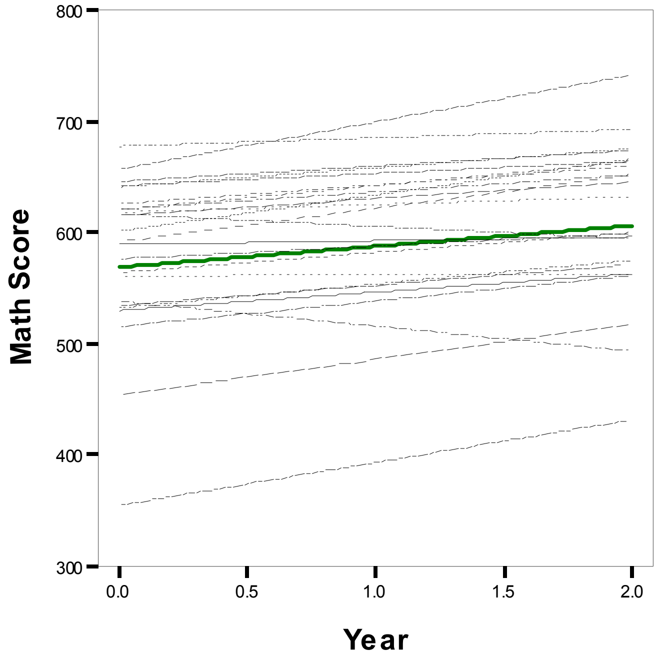
Figures 16a and 16b indicate that there is much more variability in the intercepts (2002 grade 8 test scores) and slopes (growth rates) of non-IB students than for IB students. The red line in Figure 16a and the green line in Figure 16b represent the average regression lines for the two groups of students. Inspection of the two figures demonstrates the significantly higher mean initial status (mathematics score in Year 0) and the steeper growth trajectory in the IB sample.

Figure 16a
Ordinary Least Squares Mathematics Growth Trajectories
Grades 8-10, 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 16b
Ordinary Least Squares Mathematics Growth Trajectories
Grades 8-10, 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 17a and 17b 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 mathematics scores. Note the much greater variability, indicated by the confidence interval bands around the two mean regression lines, in the growth trajectories of the non-IB students.

Figure 17a
Mean OLS Regression Line for Mathematics Growth
Grades 8-10, 2002-2004 for a Sample of IB Students

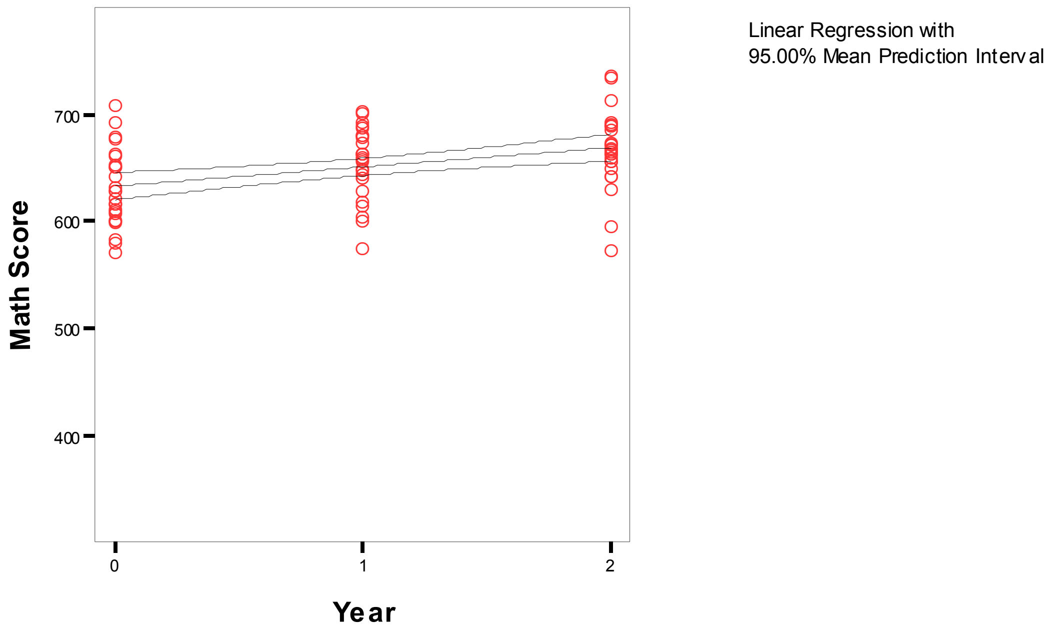
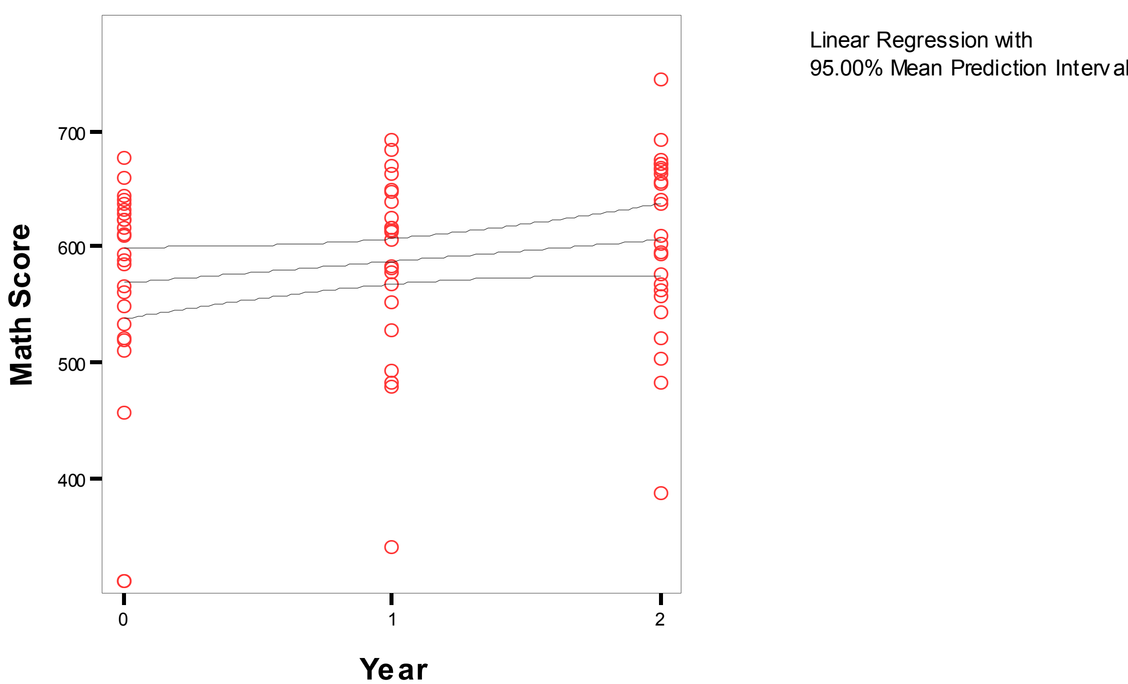


Figure 17b
Mean OLS Regression Line for Mathematics Growth
Grades 8-10, 2002-2004 for a Sample of Non-IB Students



5.2.2 Results from the Multiple Regression: Exploratory Predictive Models

In order to identify factors that are potentially predictive of student achievement and growth in mathematics, bivariate correlations and stepwise linear regression models were estimated prior to the hierarchical linear modeling. The bivariate correlations indicate that the strongest correlates to the outcome, mathematics achievement in tenth grade, are reading scores in grade 8 ($r = .71$) and in grade 10 ($r = .74$) and prior mathematics achievement in grade 8 ($r = 0.87$. The strongest correlate to average gain in achievement from eighth through tenth grades is reading achievement in grade 10, but that correlation is very weak, $r = .24$, $p < .001$). Prior math achievement (grade 8) appears even more weakly related to average growth in math from grade 8 to grade 10 ($r = -.13$, $p < .001$). Variables such as gender, minority status, language background, and poverty status do not appear to be significantly related to either math achievement or growth in math achievement.

The preliminary multiple regression analyses indicate that the best predictors of math achievement in tenth grade and average gain in achievement are grade 8 math score (math initial status), grade 10 reading score, and gender. The next section describes the results of the hierarchical linear models (HLM) of mathematics achievement and growth during high school in District 20.

5.3 Results from the HLM Estimation of the Effects of IB Participation on Mathematics Growth in High School

The OLS fits to the raw achievement data discussed in § 5.1.3 and § 5.2.1 indicate sufficient variability to warrant more sophisticated modeling. Hierarchical Linear Models (HLM) was used to examine the effects of IB participation on mathematics growth and outcomes. The results of the HLM models are described below.

The preliminary regression analyses indicate that only initial mathematics status (grade 8 math score), grade 10 reading achievement, and gender are significant predictors of mathematics outcomes and growth in District 20's high schools. The HLM analyses generally confirm these results. As will be described in § 5.3.2 and § 5.3.3, there appears to be significant main effects for both IB participation and number of years of participation in the program. However, these effects appear to be confounded by general academic achievement.

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 10 2004 CSAP mathematics assessment.

5.3.1 Unconditional Means and Growth Models

As described in § 4.3.1 and § 4.3.2 of the middle school analyses, an unconditional means model and unconditional growth model are first estimated in order to, respectively, (1) partition the outcome variation and to (1) indicate 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. By partitioning the variance into within-persons (growth) and between-persons (individual factors, including those relating to selection effects) without controlling for additional person-level variables, the unconditional model provides the baseline for evaluating the adequacy of subsequent models.

The Level 1 and Level 2 submodel specifications are presented in § 4.3.1 and § 4.3.2, and are not repeated here, but the results of the of the unconditional means and growth models for the 2002-2004 grade 8-10 cohort are provided in Table 10.

Table 10
Comparison of Unconditional Means and Growth Models
of Mathematics Performance and Growth
Grades 8-10, 2002-2004

		Parameter	Unconditional Means Model	Unconditional Growth Model
Fixed Effects				
Initial status, π_{0i}				
Mean grade 5 math score		β_{00}	599.07*** (1.83)	587.24*** (1.82)
Rate of change, π_{1i}				
Mean student growth rate		β_{10}		11.83*** (0. 50)
Variance Components				
Level 1	Within-persons	e_{ti}	577.95	394.38
Level 2	For initial status	r_{0i}	3024.52***	2880.91***
	For rate of change	r_{1i}		42.62***
	Correlation	r_{01}		0.23
Goodness-of-fit				
Deviance			29243.44	28680.81
χ^2 statistic				562.63***

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, π_{0i} (person i's average mathematics scores regardless of time). Thus, the intercept represents the average of the grade 8 – 10 math scores averaged across all students. 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 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 (ρ) obtained from the unconditional means model is:

$\rho = r_{0i} / (r_{0i} + e_{ti}i)$
$\rho = 3024.52 / (3024.52 + 577.95) = 0.8396$

Thus, approximately 84 percent of the total variation in mathematics 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 mean student scores at the initial measurement (i.e., grade 8). Thus, students average about 587 on the initial grade 8 mathematics assessment (β_{00}) and grow at an average rate of about 12 scale score points per year (β_{10}) from eighth 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 (the growth trajectory) by 32 percent, signifying that there is true growth in mathematics between eighth and tenth grades:

$[(577.95 - 394.38) / 577.95] = .318.$

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 mathematics status) and rate of change (mathematics growth). Several conditional models are estimated, but only the two final models are described. Exploratory analyses and initial conditional models indicate that gender and reading levels in grade 8 are related to math achievement in grade 8, but none of the hypothesized covariates (e.g., initial status, gender) significantly predict mathematics <u>growth</u> once IB variables are accounted for. The model described in §5.3.2 evaluates the effect of high school IB participation on mathematics achievement and growth from eighth through tenth grades. This model conceptualizes IB as a "treatment" effect. The model described in §5.3.3 evaluates the effect of number of years of middle and high school participation in the IB programs on mathematics achievement and growth from eighth through tenth grades.

5.3.2 Final Conditional Model: Effect of IB Participation on Mathematics Achievement

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 sufficiently "explain" all the observed variation in mathematics growth. Several exploratory models estimated in this iterative process indicate the following:

- Initial math status (achievement in Grade 8) does not appear to be related to growth in math from grade 8 to 10;
- Reading performance in grade 8 does not appear to be related to growth in math from grade 8 to 10; and
- IB participation in grade 8 does not appear to be related to growth in math from grade 8 to 10. Note that IB participation in grade 8 does not imply participation in grade 10. Only 62 of the 101 (61%) grade 8 IB students were still in the IB program in grade 10[12].

The results for the best fitting conditional model of effects of IB participation on growth in math achievement are provided in Table 11. This model examines the effects of gender and grade 8 reading performance on grade 8 mathematics achievement and effects of IB participation in high school on mathematics growth rate (slope) during high school without controlling for other student characteristics. The Level 1 and 2 submodels are:

LEVEL 1 MODEL

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

LEVEL 2 MODEL

$$\pi_{0i} = \beta_{00} + \beta_{01}(\text{GENDER}) + \beta_{02}(\text{READ02}) + r_{0i}$$

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

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

$$\text{MATHEMATICS}_{ti} = \beta_{00} + \beta_{01}(\text{GENDER}) + \beta_{02}(\text{READ02}) + \beta_{10}(\text{YEAR}) + \beta_{11}(\text{IB*YEAR})) + r_{0i} + r_{1i}(\text{YEAR}_{ti}) + e_{ti}$$

where,

MATHEMATICS_{ti} is the outcome variable, mathematics score in grade 10

π_{0i} = Initial mathematics status of person *i*, that is, the expected outcome for that student in the spring of grade 8 (when Year = 0)

π_{1i} = Rate of change in mathematics (growth rate) for person *i*

[12] And who also met study requirements.

YEAR is 0 at spring 2002, 1 at spring 2003 and 2 at Spring 2004 (grades 8, 9, & 10)
β_{00} = Estimated mean intercept, or initial mathematics status of students at year 0
β_{01} = Mean effect of gender on the intercept (initial mathematics status)
β_{02} = Mean effect of reading performance on the intercept (initial mathematics status)
β_{10} = Mean academic year growth rate in mathematics
β_{11} = Mean effect of IB on mathematics growth rate
IB = 0 if student was not in the IB program in 2003-04, and 1 if student was
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 individuals in the population
r_{1i} = Level 2 residual variance in true rate of change across all individuals in the population (between-persons deviation)

The fixed and random effects from this model are provided in Table 11. This model attempts to control for possible selection effects by controlling for the effect of reading ability on initial math status (grade 8 math score)[13]. Reading performance, as well as prior performance in other subjects, may be interpreted as incorporating the effects of the numerous personal and contextual factors that influence academic achievement and gains in achievement, including prior learning, opportunities to learn, extrinsic and intrinsic motivation, parental desires, home and community environment, etc. By attempting to control for reading achievement, we are making an effort to include the effects of other unmeasurable and unknowable factors.

[13] Another model attempted to evaluate selection effects on mathematics growth by estimating the effects of grade 8 reading level on math growth rates, but the effect was close to zero and not significant.

Table 11
Effects of Reading Performance and IB Participation on
Mathematics Performance and Growth, Grades 8-10, 2002-2004

		Parameter	Coefficient
Fixed Effects			
Initial status, π_{0i}			
Mean grade 5 math score)		β_{00}	594.84***
			(1.83)
Gender effect		β_{01}	-15.45
			(2.39)
Mean reading effect		β_{02}	0.89***
			(0.03)
Rate of change, π_{1i}			
Mean growth rate for non-IB Students		β_{10}	11.41***
			(0. 52)
IB effect on growth rate (IB – non-IB gap)		β_{11}	6.59***
			(1.29)
Variance Components			
Level 1	Within-persons	e_{ti}	394.37
Level 2	For initial status	r_{0i}	1163.59***
	For rate of change	r_{1i}	41.01***
Goodness-of-fit			
Deviance			27870.11
χ^2 statistic			40.04***

Numbers in parentheses represent standard errors. *** $p < .001$

The variance decomposition provided in the bottom panel of Table 11 indicates that students' gender and reading level in grade 8 account for 60 percent of the residual variance in grade 8 mathematics score.[14] In other words, about 60 percent of the variability in students' grade 8 math performance can be "explained" by gender and their reading ability at the time of the assessment. However, it appears that the addition of IB participation explains only about four percent of the residual between-persons variance in rate of change.[15]

[14] The percent reduction in residual variance for the intercept that is between persons is estimated as percentage change from the unconditional model: [(2880.91 – 1163.59) / 2880.91] = 0.5961.
[15] I.e., [(42.62 – 41.01) / 42.62] = 0.0378.

On average, males who were average readers on the grade 8 assessment scored about 595 points (β_{00}) on the 2002 grade 8 CSAP mathematics assessment. The gender gap is -15 points, indicating that, on average, females score about 15 points lower than males on the grade 8 math assessment. Grade 8 reading performance is positively related to math achievement: in eighth grade, a one point increase in the reading score is associated with a 0.89 point increase on the mathematics assessment.

The most significant result, from a theoretical standpoint is that participation in the diploma IB program appears to have a statistically significant positive impact on growth in mathematics achievement. The average growth rate from eighth to tenth grade is about 11 points, with IB students gaining about six and a-half points per year on their non-IB counterparts.

Three striking results are shown in Table 11:

1. **IB students in the diploma program demonstrate significant average annual gains over their non-IB counterparts, although comparison of residual variance indicates almost negligible reduction in the amount of slope variance when IB is included in the model.**

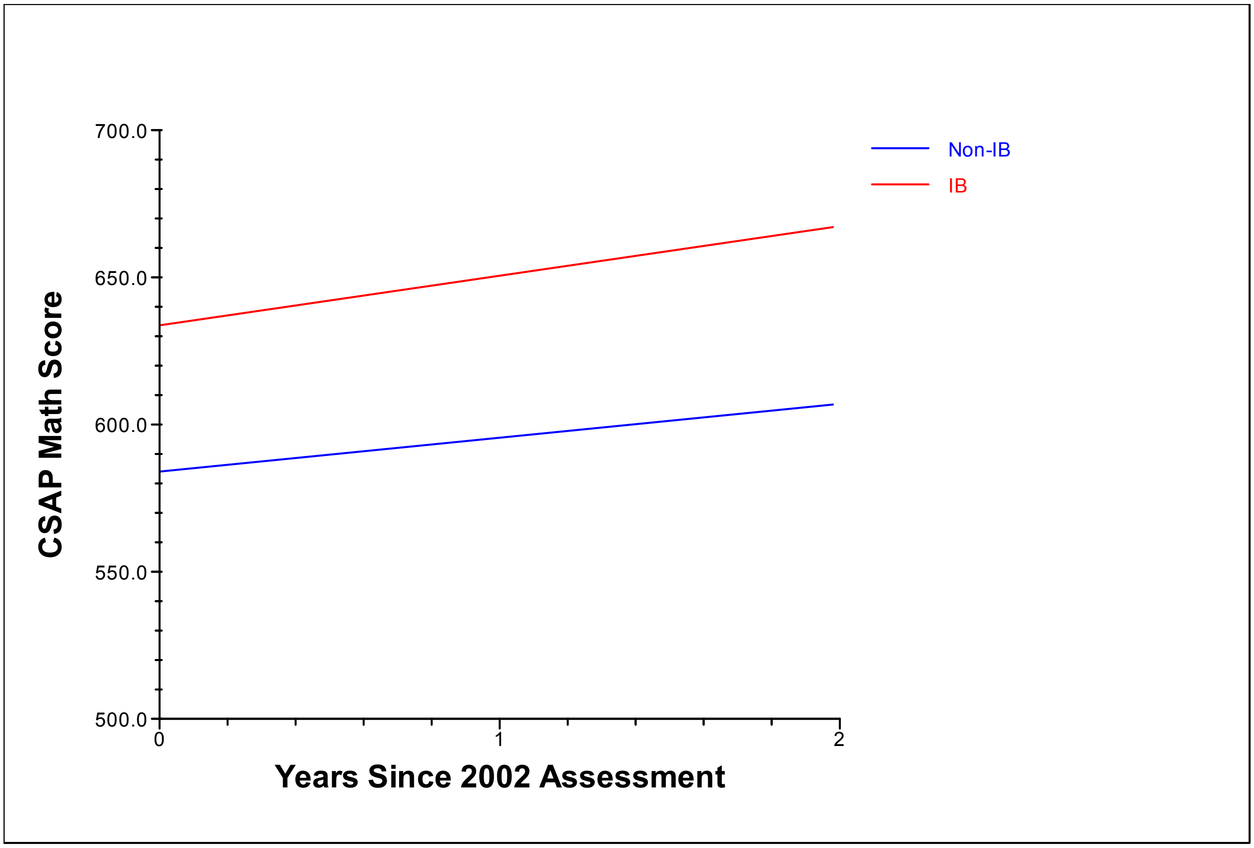
2. **Students' mathematics performance in grade 8 is strongly related to their reading ability (as assessed on the grade 8 CSAP). For every extra score point on the reading assessment, students, on average, score about nine-tenths (0.89) of a point higher on the grade 8 math assessment (β_{01}).**

3. **Males out-score females by 15 points, on average, on the grade 8 mathematics assessment.**

Figure 18 illustrates the differences in the mathematics achievement and gains of IB and non-IB students in this cohort. This figure clearly indicates that not only do IB students score higher than non-IB students, their average rate of gain during high school is also greater than that of other students. This is evident from the significantly steeper slope of the IB students.

Another model was estimated, in which grade 10 reading performance (2004) was included in the slope submodel. The results of this model indicate that reading performance on the tenth grade assessment (2004) <u>does</u> appear to be related to how much gain students demonstrate in years <u>prior to</u> the grade 10 assessment. Furthermore, when reading is included along with IB in the slope submodel, the level and significance of the IB effect drop substantially. When 2004 reading score is included, the IB effect drops from 6.04 points to 2.3 points, and significance level drops from $p < .001$ to $p = .10$, even though the reading effect appears inconsequential – a one-point increase in the 2004 reading score is associated with *one-tenth* of a point increase in growth rate. Thus, it appears that the effects of IB and general academic achievement (indicated by reading level) may be confounded. In sum, these results indicate that IB students generally score higher than non-IB students and IB students' growth rate in mathematics tend to increase at a faster rate than that of non-IB students, but some of these differences may be partially due to selection effects.

Figure 18
Mathematics Growth Trajectories of IB and Non-IB Students
Grade 8-10 Cohort, 2002-2004



<u>5.3.3 Final Model: Effects of Number of Years of IB Participation on Mathematics Achievement</u>

The model described in §5.3.2 evaluates the effect of high school IB participation on mathematics achievement and growth from eighth through tenth grades. The model described below evaluates the effect of number of years of middle and high school participation in the IB programs on mathematics achievement and growth from eighth through tenth grades.

Several iterative models, similar to those described in the above section, were estimated, but the IB variable in these models is number of years in the middle years and diploma IB programs, rather than IB participation in high school. The final model illustrated below examines the effects of gender and grade 8 reading performance on grade 8 mathematics achievement and effects of number of years of IB participation on mathematics growth rate (slope) during high school without controlling for other student characteristics. The Level 1 and 2 submodels are:

LEVEL 1 MODEL

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

LEVEL 2 MODEL

$$\pi_{0i} = \beta_{00} + \beta_{01}\,(\text{GENDER}) + \beta_{02}\,(\text{READ02}) + r_{0i}$$

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

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

$$\text{MATHEMATICS}_{ti} = \beta_{00} + \beta_{01}\,(\text{GENDER}) + \beta_{02}\,(\text{READ02}) + \beta_{10}\,(\text{YEAR}) + \beta_{11}\,(\text{YRS\_IB*YEAR})) + r_{0i} + r_{1i}\,(\text{YEAR}_{ti}) + e_{ti}$$

where,

MATHEMATICS_{ti} is the outcome variable, mathematics score in grade 10

π_{0i} = Initial mathematics status of person *i*, that is, the expected outcome for that student in the spring of grade 8 (when Year = 0)
π_{1i} = Rate of change in mathematics (growth rate) for person *i*
YEAR is 0 at spring 2002, 1 at spring 2003 and 2 at Spring 2004 (grades 8, 9, & 10)
β_{00} = Estimated mean intercept, or initial mathematics status of students at year 0
β_{01} = Mean effect of gender on the intercept (initial mathematics status)
β_{02} = Mean effect of reading performance on the intercept (initial mathematics status)
β_{10} = Mean academic year growth rate in mathematics
β_{11} = Mean effect of years in the IB program on mathematics growth rate
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 individuals in the population
r_{1i} = Level 2 residual variance in true rate of change across all individuals in the population (between-persons deviation)

The fixed and random effects from this model are provided in Table 12. This model attempts to control for possible selection effects by controlling for the effect of reading ability on initial math status (grade 8 math score)[16]. Reading performance, as well as prior performance in other subjects may reflect effects of the numerous personal and contextual factors that influence academic achievement and gains in achievement. By attempting to control for reading achievement, we are making an effort to include the effects of other unmeasurable and unknowable factors, including selection effects.

[16] Another model attempted to evaluate selection effects on mathematics growth by estimating the effects of grade 8 reading level on math growth rates, but the effect was close to zero and not significant.

Table 12
Effects of Reading Performance and Number of Years of IB Participation on Mathematics Performance and Growth, Grades 810, 2002-2004

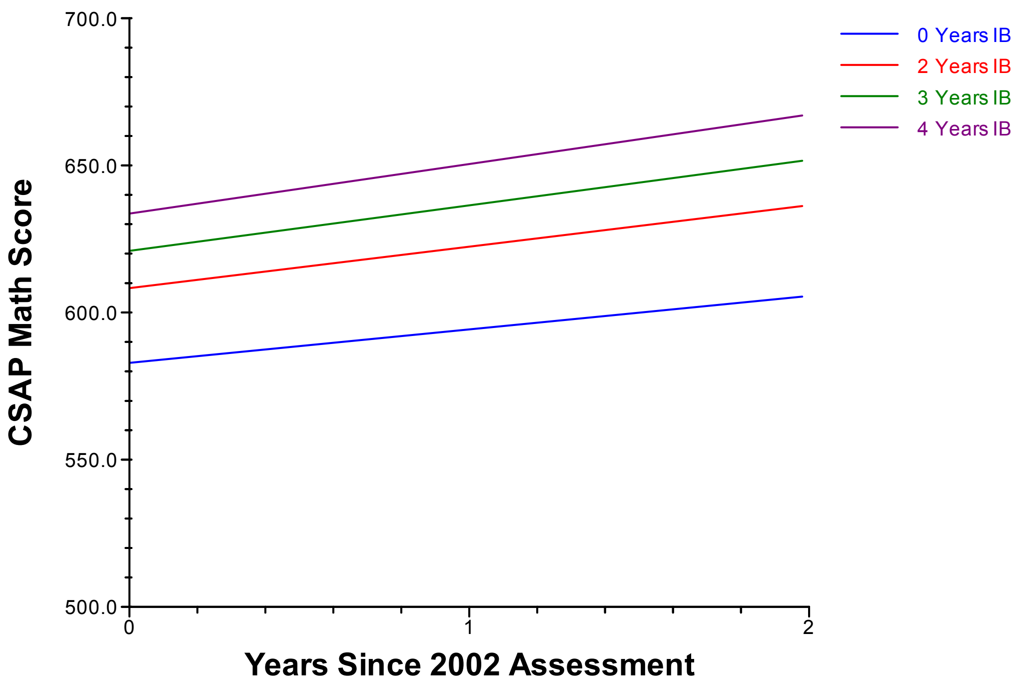
		Parameter	Coefficient
Fixed Effects			
Initial status, π_{0i}			
Mean grade 8 math score		β_{00}	594.84*** (1.83)
Gender effect		β_{01}	-15.45 (2.39)
Mean reading effect		β_{02}	0.89*** (0.03)
Rate of change, π_{1i}			
Mean growth rate for non-IB Students		β_{10}	11.27*** (0. 53)
Years IB effect on growth rate (IB – non-IB gap)		β_{11}	1.64*** (1.29)
Variance Components			
Level 1	Within-persons	e_{ti}	394.37
Level 2	For initial status	r_{0i}	1163.54***
	For rate of change	r_{1i}	40.75***
Goodness-of-fit			
Deviance			27869.04
χ^2 statistic			12.45**

Numbers in parentheses represent standard errors. *** $p < .001$ ** $p = .001$

The results of this model are extremely similar to those of the previous model of IB participation (see Table 11). As expected, the mean, intercept, slope, and fixed effects for gender and 2002 reading performance are the same. The results of this model also indicate that number of years in the IB program from grade 7 through grade 10 has a small, but significant, positive effect on growth in mathematics achievement. On average, one year of participation results in a 1.6 point increase in growth rate (p <.001). Compared to the average growth rate of non-IB students (11.3 points), the average growth rate of IB students is 12.9 points per year. This implies, for example, that students who participate in IB for four years during this time period should average about 6 points higher on the grade 10 math assessment than non-IB students.

Figure 19 illustrates the effects of IB participation and number of years of participation in the IB program. Students in this cohort may be described by four categories: never IB between 2001 and 2004 (0 years); two years in IB; three years in IB; or four years in IB. No students in this analytical cohort participated either one year or five years.

Figure 19
Mathematics Growth Trajectories of High School Students by
Number of Years in the IB Program
Grade 8-10 Cohort, 2002-2004



Note the substantially lower average performance and flatter slope of students never in the IB program. Also note the fan-like pattern, which indicates that average performance is more tightly clustered at the intercept than in year 2. This indicates that more years in the IB program are associated with greater increases (gains) in student performance in mathematics.

Another model was estimated, in which tenth grade reading performance (2004) was included in the slope submodel. The results of this model indicate that reading performance on the tenth grade assessment (2004) appears to be related to how much gain students demonstrate in years prior to the grade 10 assessment. Furthermore, when both years of IB participation and reading are included in the slope submodel the level and significance of the years in IB effect drop substantially. When 2004 reading score is included, the IB effect drops from 1.64 points to

0.57 points, and significance level drops from $p < .001$ to $p = .10$, even though the reading effect appears inconsequential – a one-point increase in the 2004 reading score is associated with *one-tenth* of a point increase in growth rate. Thus, it appears that the effects of number of years in the IB program and general academic achievement (indicated by reading level) may be confounded. In sum, these results indicate that IB students generally score higher than non-IB students and IB students' growth rate in mathematics tend to increase at a faster rate with additional years of IB participation. However, some of these differences may be partially due to selection effects.

5.4 Summary of Results for the 2002-2004 Grade 8-10 Cohort

In summary, the analyses reported above indicate that IB students score significantly higher on the CSAP mathematics assessments of grades 8 through 10. The growth rate of these IB students also appears to be somewhat significantly higher than that of non-IB students in general and much higher than that of non-IB students at Rampart High School. Furthermore, additional years of participation in the IB program appear to be associated with greater gains in mathematics achievement. Results of the analyses also indicate that males out-score females by an average of 15 points on the grade 8 math assessment and students' mathematics performance in grade 8 is strongly related to their reading ability.

6. Conclusions

Exploratory models indicate that unlike reading performance and growth, initial math status does not appear to be related to growth in math during middle or high school. Further, reading performance in grades 5 and 8 does not appear to be related to growth in math from grades 5 to 8 or 8 to 10. The most striking results from these analyses are:

1. Although average grade 5-8 mathematics scores are higher for students in the middle years IB program, their growth rate appears lower than for non-IB students. Thus, the middle school program does not appear to provide an obvious acceleration to students' learning growth in mathematics.
2. Students who are in the diploma IB program in grade 10 score significantly higher on the CSAP mathematics assessments of grades 8 through 10.
3. IB students in the diploma program demonstrate significant average annual gains over their non-IB counterparts.
4. Additional years of participation in the IB program during middle and high school appear to be associated with greater gains in mathematics achievement in high school.
5. Students' mathematics performance in grade 5 is strongly related to their reading ability in grade 5.
6. Students' mathematics performance in grade 8 is strongly related to their reading ability in grade 8.
7. Males out-score females by 15 points, on average, on the grade 8 mathematics assessment.

Although this study suggests that IB participation has a positive impact on student achievement and gains in mathematics during middle and high school, the main effects are relatively small. However, we do not have sufficient data to adequately explain selection effects (e.g., family income and educational level, student's interest in mathematics, motivation, parental encouragement to enroll in IB). We also cannot know how these IB students would have performed without the IB program. Such an evaluation can only be made through a randomized experimental design, in which students are randomly assigned to the IB or regular education program, irrespective of students' prior achievement, students' desires, parental desires, student demographics, geography, etc.

Another avenue for future research is to conduct a qualitative study of District Twenty's IB program in an attempt to disentangle effects that may be attributable to the program itself, student background and self-selection, how the program is administered in the different sites, teacher effects, etc. Such a study should help the district identify what aspects of the program appear to be critical for improving student learning and achievement with the goal of applying new strategies or strengthening the instructional practice and/or organizational conditions in other classrooms to help teachers and schools more fully support student academic learning and achievement. Such a study also should help identify areas where the program, or its administration, appear to be breaking down.