

IB PRIMARY YEARS PROGRAM EVALUATION

Prepared for Pasadena Unified School District

May 2019



In the following report, Hanover Research evaluates the effects of the International Baccalaureate Primary Years Program in Pasadena Unified School District. This report focuses on academic and behavioral outcomes for students who participated in the program in the 2008-09 through 2017-18 school years.

TABLE OF CONTENTS

Executive Summary.....	3
INTRODUCTION	3
KEY FINDINGS.....	4
RECOMMENDATIONS	4
Section I: Data Overview and Methodology	5
DATA OVERVIEW	5
METHODOLOGY	9
Section II: Effects of IB PYP	14
EFFECTS FOR NON-IB MIDDLE SCHOOL STUDENTS	14
EFFECTS FOR BLAIR HS MIDDLE SCHOOL STUDENTS	18

EXECUTIVE SUMMARY

INTRODUCTION

In this report, Hanover Research (Hanover) evaluates the effects of the International Baccalaureate Primary Years Program (IB PYP) in Pasadena Unified School District (PUSD). Since 2002-03, Willard Elementary School (Willard ES) has offered the IB PYP to all students from Kindergarten through Grade 5. PUSD is interested in evaluating the impact of the program on the participants' academic and behavioral outcomes in later grade levels compared to the outcomes of students who attended other elementary schools in the district (i.e., non-participating students).

In this analysis, Hanover focuses on students who 1) **entered PUSD in Kindergarten or Grade 1**; 2) **enrolled in a PUSD elementary school for at least five years**;¹ and 3) **enrolled in a PUSD middle school for at least one year from Grades 6 to 8**. Students who enrolled at Willard ES for at least five years are considered IB PYP participants, while students who enrolled in another elementary school are classified as eligible non-participants. Limiting the sample this way allows Hanover to observe student characteristics before the impact of IB PYP participation and after the program.²

Hanover uses the propensity score matching (PSM) technique to construct a control group and compare program participants to *matched*, or similar non-participants based on their pre-program demographic characteristics (i.e., those who are demographically similar). In this report, Hanover focuses on middle school academic and behavioral outcomes.³ The analysis is conducted separately for students who attend a non-IB middle school and for those students who attend Blair High School from Grades 6 to 8, which is another PUSD school offering an IB program.

This report comprises two sections:

- **Section I** describes the **data and methodology** used in the analysis.
- **Section II** presents the comparison of **middle school outcomes** for participants and *matched similar* non-participants. The comparison group is constructed using the propensity score matching technique.

¹ For students who have repeated grade levels, Hanover only keeps the earliest record.

² There are no pre-program academic outcomes (i.e., outcomes prior to students' enrollment into Kindergarten or Grade 1) available for any of the students in the data. Thus, Hanover uses Grade 1 demographic characteristics to construct a comparison group on the assumption that this set of parameters accounts for all potential pre-program differences between IB PYP participants and non-participants.

³ The current analysis focuses on outcomes in Grade 6 through 8. We could not conduct an evaluation of Grade 11 outcomes because the three cohorts of students with Grade 11 outcomes are not observed in the provided data before Grade 2. This does not allow us to construct a comparison group of similar non-participants based on early characteristics. Additionally, due to the lack for data in early grade levels, Hanover could not determine the number of years impacted by the program or enrolled in the district.

KEY FINDINGS

- **IB PYP participants who attend a non-IB middle school perform significantly better on SBAC ELA assessments than matched non-participants.** In Grades 7 and 8, the average scaled scores for participants are 28-30 points higher, while the probabilities of meeting or exceeding standards are seven to nine percentage points higher than for matched non-participants.
 - **In addition to the overall outcomes, program participants also outperform their counterparts on the Research/Inquiry Claim of the ELA assessment.** Specifically, IB PYP participants are 15 percentage points more likely to exceed the standard in Grade 7 and 13 percentage points more likely to meet the standard in Grade 8.
 - At the same time, Hanover does not detect significant impacts of the program on SBAC Math outcomes and overall GPA.
 - **English Learners (EL) participating in IB PYP are more likely to be reclassified by Grade 8 than similar non-participating students.** The reclassification rate for EL participants is 16 percentage points higher than for non-participants.
 - **The program has a small positive effect on student attendance in Grades 6 and 7.** The average attendance rate for IB PYP participants is 0.6-0.9 percentage points higher in the first two years after the end of the program.
- **The program also has a significant positive effect on ELA outcomes for IB PYP participants who also attend an IB program during their middle school years in Blair HS.** For example, in Grade 7, the probability of meeting SBAC ELA standard for participants is 31 percentage points higher, while the probability of exceeding the standard on the Research/Inquiry Claim of the ELA assessment is 18 percentage points higher. At the same time, the average attendance rate for this group of participants is slightly lower than for similar non-participants.

RECOMMENDATIONS

- **PUSD may promote the IB PYP to prospective families as a program that develops students' language skills, which is especially relevant for ELs.** This analysis shows that IB PYP participants outperform similar students from other elementary schools in the district on middle school SBAC assessments in ELA. At the same time, ELs participating in the IB program in Willard ES are more likely to be reclassified by Grade 8 than students attending other schools in the district.
- **Hanover recommends PUSD run a more comprehensive analysis of outcomes for students attending Willard ES in elementary school and Blair HS in later grade levels within the next several years.** The analysis indicates that continuing with IB in middle school only recently became a popular option among Willard ES graduates. If this trend continues, a future evaluation would be able to use a larger sample size allowing for more robust estimates.

SECTION I: DATA OVERVIEW AND METHODOLOGY

In this section, Hanover describes the data provided by PUSD for this analysis, as well as the methodology employed to address the district’s research questions.

DATA OVERVIEW

PUSD provided Hanover with demographic and academic data from 2008-09 through 2018-19. The data includes the following components:

- **Enrollment and Program Participation** – enrollment information (e.g., student ID, year of enrollment, school of enrollment, grade level, etc.) for students enrolled in PUSD schools between 2008-09 and 2018-19. Hanover uses enrollment in Willard ES in Grades K-5 as an indicator of participation in the IB PYP.
- **Demographic Characteristics** – students’ gender, race/ethnicity, EL status, special education status, socioeconomic disadvantaged status in each school year when a given student is enrolled in PUSD.
- **Academic Performance** – Smarter Balanced Assessment (SBAC) ELA and Math outcomes and grade point averages (GPA) in each year when a given student is enrolled in PUSD.
- **Behavioral Performance** – information on disciplinary incidents and students’ school attendance in each year when a given student is enrolled in PUSD.

Figure 1.1 presents the variables of interest in the IB PYP evaluation analysis and the assumptions made with regard to each data point.

Figure 1.1: Descriptions and Data Assumptions of the Variables of Interest

VARIABLE	DESCRIPTION	DATA ASSUMPTION
Enrollment and Program Participation		
Enrollment Identifiers	Students’ enrollment records, including grade level and school of enrollment in each school year.	<p>131 students in the data (0.3 percent of all) are enrolled in multiple schools within the same year. Hanover keeps the observations labelled with “Primary Enrollment” in the Enrollment Status Variable.</p> <p>Additionally, if a student repeats a grade level, Hanover keeps the earliest attempt for each student.</p> <p>In the final analytic sample, Hanover aggregates some of the schools together for the purposes of regression analysis. Specifically, Sierra Madre MS and CIS Academy are grouped with Wilson MS, McKinley is grouped with Washington MS, and Eliot MS is grouped with Marshall Fundamental.</p>
Demographic Characteristics		
Gender	Female, male.	--

VARIABLE	DESCRIPTION	DATA ASSUMPTION
Race/Ethnicity	American Indian/Alaska Native, Black/African American, Asian, Filipino, Hispanic, Multiple, Native Hawaiian/Pacific Islander, White, Not Reported, and Missing.	Due to small sample size, Hanover aggregates Missing and Multiple under Multiple or Unknown category. Asian and Native Hawaiian/Pacific Islander are aggregated under Asian/Pacific Islander category.
English Learner	EL, EO, FEP, RFEP	<p>The data for EL status appears inconsistent across years in the CALPADS Enrollment files. Hanover uses the LF variable from Aeries Database from 2008-09 through 2017-18 and uses the ELL status variable from the CALPADS Enrollment files in 2018-19.</p> <p>Students who are labelled with RFEP, EO, or FEP in one year and labelled with EL in at least one of the previous years are assumed to be reclassified. For each EL student, Hanover also calculates a variable - Years in EL - which is a total number of years that a student spent in PUSD EL program before reclassification.</p>
Special Education Status	Y, N.	--
Economic Disadvantage Status	Y, N.	The values of this variable are drastically different in 2008-09 as compared to later years. Hanover does not use these values in 2008-09 in the analysis.
Academic Performance		
Smarter Balanced Assessment (SBAC)	SBAC Achievement Levels and Scaled Scores for ELA and Math in Grades 3-8 and Grade 11.	In addition to Achievement Levels and Scaled Scores, Hanover also looks at the ELA Research/Inquiry Claim outcomes.
GPA	Students' grade point average in each school year.	Hanover uses the variable GPN from Aeries Database which has a range 0-4 assuming it to be unweighted GPA. Hanover also assumes that zero values in GPA variable stand for missing GPA.
Behavioral Performance		
Attendance Data	Students' days enrolled and days present in each school year.	<p>For students who have duplicate rows at student-year level in the Aeries Database, Hanover keeps the record with the maximum value of days enrolled.</p> <p>Hanover defines attendance rate as a ratio of days present to days enrolled.</p>
Incidents Data	Incident records in each school year.	For each student, Hanover calculates the total number of incidents and total number of incidents resulting in out-of-school suspension. Due to small variability in both variables, Hanover converts them to a binary format, i.e. 0 for no incident or OSS and 1 for at least one incident or OSS.

To capture the effect of the IB PYP, Hanover focuses its analysis on students whose characteristics and outcomes are **observable in the data before and after the program**

impact for valid comparison between program participants and non-participants. **Figure 1.2** shows the five cohorts of PUSD students who satisfy these criteria. Cells highlighted in red indicate years and grade levels in which students in Willard ES participate in IB PYP. Cells highlighted in blue indicate the middle school years.

Note that to increase the available sample size Hanover includes in the analysis, Cohort 1 students who are observed for the first time in the data in 2008-09 while they are in Grade 1. This assumes that students’ pre-program characteristics in Grade 1 are not drastically different from their characteristics in Grade K, which is unobservable in the data (i.e., 2007-08 school year). This assumption is reasonable in the research setting because Hanover only considers students’ demographic characteristics, which are unlikely to be affected by one year of participation or non-participation in IB PYP.

Figure 1.2: Student Cohorts of Interest

GRADE	K	GRADE 1	GRADE 2	GRADE 3	GRADE 4	GRADE 5	GRADE 6	GRADE 7	GRADE 8
2008-09	Cohort 2	Cohort 1							
2009-10	Cohort 3	Coh. 2	Coh. 1						
2010-11	Cohort 4	Coh. 3	Coh. 2	Coh. 1					
2011-12	Cohort 5	Coh. 4	Coh. 3	Coh. 2	Coh. 1				
2012-13		Coh. 5	Coh. 4	Coh. 3	Coh. 2	Coh. 1			
2013-14			Coh. 5	Coh. 4	Coh. 3	Coh. 2	Coh. 1		
2014-15				Coh. 5	Coh. 4	Coh. 3	Coh. 2	Coh. 1	
2015-16					Coh. 5	Coh. 4	Coh. 3	Coh. 2	Coh. 1
2016-17						Coh. 5	Coh. 4	Coh. 3	Coh. 2
2017-18							Coh. 5	Coh. 4	Coh. 3

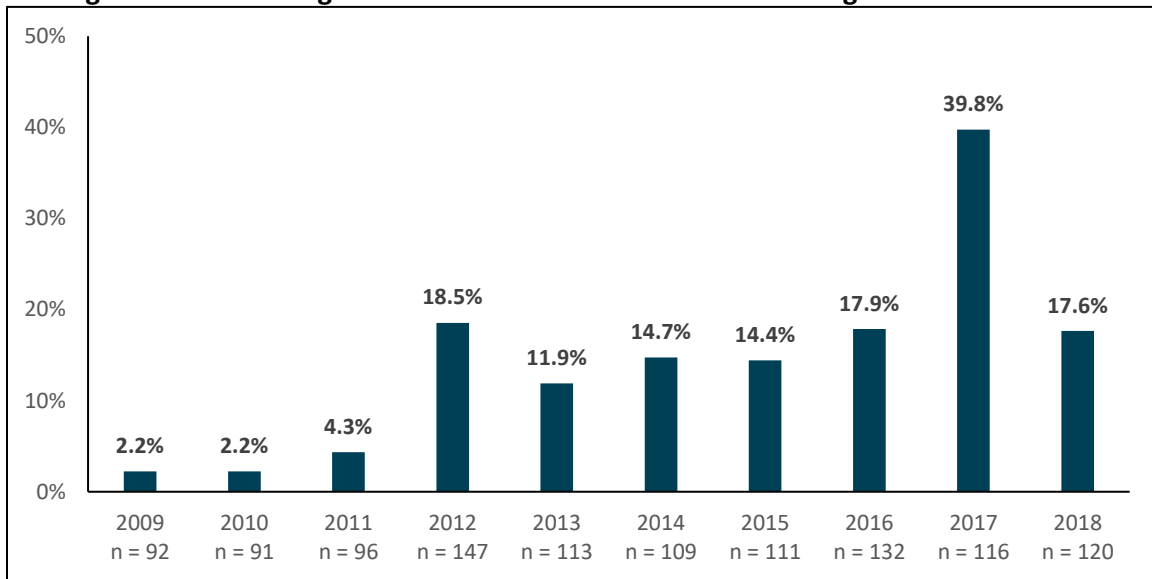
As discussed above, Hanover only focuses on students who 1) **entered PUSD in Kindergarten or Grade 1**; 2) **enrolled in a PUSD elementary school for at least five years**; and 3) **enrolled in a PUSD middle school for at least one year from Grades 6 to 8**. Among these students, those who enrolled in Willard ES for at least five years from Kindergarten or Grade 1 to Grade 5 are considered **IB PYP participants** (i.e., treatment group), while those who enrolled in another elementary school are treated as **eligible non-participating students** (comparison group).

Additionally, Hanover segments the analysis by the middle school of enrollment. For the comparison between IB PYP participants and non-participants to be valid, the **middle school outcomes of interest should be measured in a similar learning environment**. However, some Willard ES students continue their education in the IB Middle Years Program (MYP) at Blair HS, while others enroll in non-IB middle schools in the district. Likewise, while some students from non-IB elementary schools continue following the non-IB route in middle school, others enroll in the IB program at Blair HS. To account for potential differences in educational setting, Hanover runs the following separate comparisons:

- Comparison of IB PYP participants and non-participants enrolling in **non-IB middle schools only**. The resulting final analytic file contains *3,918 non-participating students and 238 IB PYP participants*.
- Comparison of IB PYP participants and non-participants enrolling in **Blair HS only**.⁴ The resulting final analytic file contains *354 non-participating students and 62 IB PYP participants*.

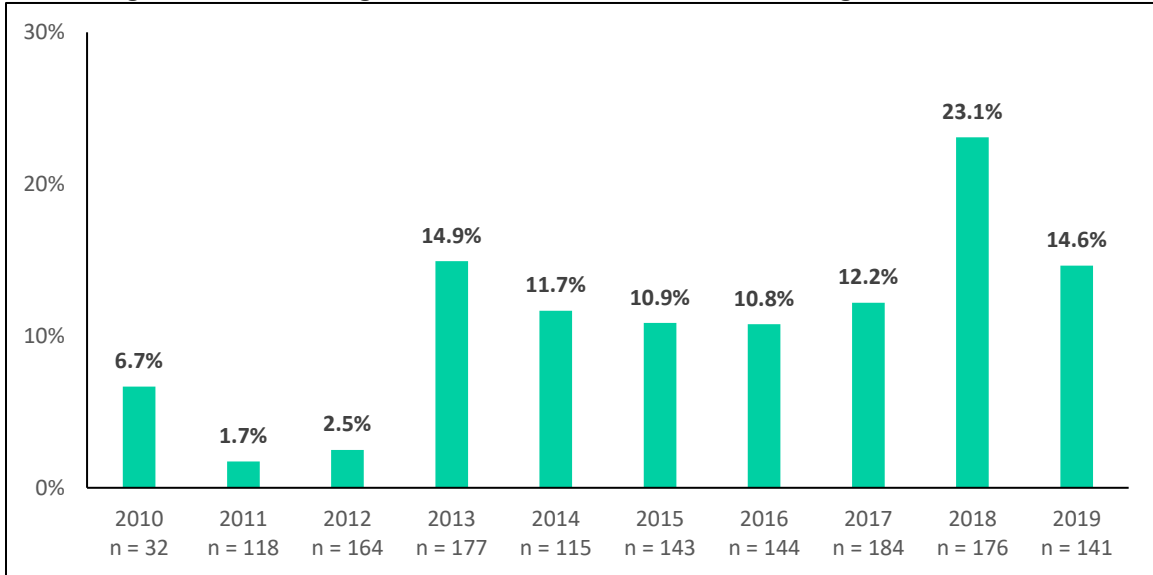
Note that the available sample size in the second comparison is significantly smaller because a relatively low percentage of Willard ES students from the cohorts observed in the data eventually enrolled in Blair HS. **Figure 1.3** shows that the percentage of Willard ES Grade 5 students who enrolled in Grade 6 in the next year was rather low in the end of the previous decade but has increased over the years to a relatively stable level – approximately 18 percent – with the exception of 2016-17, when the percentage is as high as 40 percent. Similarly, the percentage of Grade 6 incoming class in Blair HS coming from Willard ES (shown on **Figure 1.4**) followed the same pattern during this period.

Figure 1.3: Percentage of Willard ES Grade 5 Students entering Blair HS in Grade 6



⁴ In this analysis, we only include students who are enrolled in Blair HS in all of their observed middle school years (Grade 6-8) from 2013-14 through 2018-19. To be specific, for Cohorts 1 to 4, we only include students who are enrolled in Blair HS from Grades 6 to 8; while for Cohort 5, we only include students who are enrolled in Blair HS in Grades 6 and 7.

Figure 1.4: Percentage of Blair HS Grade 6 Students coming from Willard ES



METHODOLOGY

Hanover focuses on the effects of the program in Section II of this report. A critical step in evaluating the real effects of participating in the IB PYP on student outcomes is to identify an appropriate comparison group. Students with certain characteristics may be more likely to take part in the program than others, and these characteristics may be related to the performance outcomes. Particularly, the outcomes may be affected by the demographic differences between elementary school zones in PUSD. Thus, a direct comparison between IB PYP participants and non-participants may not reflect the true program effect. In the following subsection, Hanover examines the differences between program participants and eligible non-participants.

DIFFERENCES BETWEEN PARTICIPANTS AND NON-PARTICIPANTS

Figures 1.5, on the following page, displays the pre-program demographic characteristics for students who enroll in non-IB middle schools. T-tests are conducted to examine the statistical significance in the mean differences between participants and eligible non-participants. As the figure shows, there are significant differences in the demographic breakdown between the two groups. Particularly, there are more economically disadvantaged students, ELs, and Hispanic students among IB PYP participants.

Figure 1.5: Student Characteristics – IB PYP Participants and Eligible Non-Participants (Non-IB Middle School)

DEMOGRAPHIC CHARACTERISTICS	ELIGIBLE NON-PARTICIPANTS		IB PYP PARTICIPANTS		DIFFERENCE
	COUNT	MEAN/SHARE	COUNT	MEAN/SHARE	
Female	3918	47.37%	238	47.48%	-0.11%
Male	3918	52.63%	238	52.52%	0.11%
American Indian/Alaska Native	3918	0.31%	238	0.00%	0.31%
Asian/Pacific Islander	3918	3.42%	238	4.20%	-0.78%
Black/African American	3918	11.74%	238	5.46%	6.28%***
Hispanic	3918	63.99%	238	79.83%	-15.85%***
Filipino	3918	1.15%	238	0.00%	1.15%*
Multiple Races or Unknown	3918	2.45%	238	2.52%	-0.07%
White	3918	16.95%	238	7.98%	8.96%***
Gifted	3918	0.00%	238	0.00%	0.00%
Special Education	3918	8.17%	238	7.98%	0.18%
Economically Disadvantaged	3153	64.70%	178	74.16%	-9.46%***
Current English Learners (EL)	3918	45.18%	238	51.26%	-6.08%*
Fluent-English-Proficient (FEP)	3918	5.92%	238	12.61%	-6.68%***
Reclassified EL (RFEP)	3918	0.31%	238	0.00%	0.31%

Note: * p < 0.1; ** p < 0.05; *** p < 0.01. The statistically significant differences are highlighted in red. Total Number of Observations: 4,156.

Likewise, there are also some significant differences between IB PYP participants and non-participants among students who eventually enroll in IB MYP in Blair HS (shown in **Figure 1.6**), especially among EL students, RFEP students, white students, as well as Hispanic students.

Figure 1.6: Student Characteristics – IB PYP Participants and Eligible Non-Participants (Enrolled in Blair HS in Middle School)

DEMOGRAPHIC CHARACTERISTICS	ELIGIBLE NON-PARTICIPANTS		IB PYP PARTICIPANTS		DIFFERENCE
	COUNT	MEAN/SHARE	COUNT	MEAN/SHARE	
Female	354	49.15%	62	48.39%	0.77%
Male	354	50.85%	62	51.61%	-0.77%
American Indian/Alaska Native	354	1.13%	62	0.00%	1.13%
Asian/Pacific Islander	354	5.93%	62	9.68%	-3.75%
Black/African American	354	10.45%	62	8.06%	2.39%
Hispanic	354	64.41%	62	51.61%	12.79%*
Filipino	354	1.98%	62	3.23%	-1.25%
Multiple Races or Unknown	354	0.85%	62	0.00%	0.85%
White	354	15.25%	62	27.42%	-12.17%**
Gifted	354	0.00%	62	0.00%	0.00%
Special Education	354	5.08%	62	9.68%	-4.59%
Economically Disadvantaged	297	58.59%	53	47.17%	11.42%
Current English Learners (EL)	354	43.50%	62	17.74%	25.76%***
Fluent-English-Proficient (FEP)	354	7.63%	62	14.52%	-6.89%*
Reclassified EL (RFEP)	354	0.28%	62	0.00%	0.28%

Note: * p < 0.1; ** p < 0.05; *** p < 0.01. The statistically significant differences are highlighted in red. Total Number of Observations: 417.

The two groups (participants and eligible non-participants) have a different demographic breakdown. As a result, a more comparable control group is required to ensure accurate estimates of the program effects. In this evaluation, Hanover uses propensity score matching (PSM) technique to construct a control group referred to as the *matched similar non-participants*. This technique allows us to identify a group of non-participants who shared similar observable characteristics with the program participants. In the following subsection, Hanover discusses the propensity score matching technique in detail.

PROPENSITY SCORE MATCHING

Implementing PSM involves estimating a logistic regression model that predicts whether a student participated in the IB PYP using Grade 1 demographic characteristics as predictor variables.⁵ The resulting equation assigns each student a *propensity score* that represents his or her predicted probability of participating in the IB PYP. Essentially, this number measures how much the student in question resembles a typical IB PYP participant in terms of the variables provided. To obtain a sample of matched non-participants, Hanover matches the program participants to eligible non-participants for whom the model generated a similar score on a one-to-one basis.

Hanover notes that the propensity score matching technique has limitations, in that it does not take any unobservable factors into account. Thus, even though the matched non-participants are identical to participants on all observed variables, they may still differ from the participants in other unmeasured aspects that are related to program participation and student outcomes.

In this evaluation, the particular concern is the lack of pre-program academic outcomes that are reflective of students' school preparedness before they are impacted by IB PYP participation or non-participation. While some of demographic parameters, such as economic disadvantage status, can be closely correlated with expected academic performance, the matched groups of students may be still differently prepared for school.

Additionally, since PUSD has an open enrollment policy for students within the district, observing a student zoned into a non-IB elementary school attending Willard ES would be an indicator of parental influence in the student's early education that may be correlated with his or her outcomes. Unfortunately, the RS (residence school) variable in the Aeries Database is missing for all years before the 2010-11 school year, which did not allow Hanover to control for school placement effects. Therefore, the analysis is further limited by the potential effect of parental influence.

Hanover can perform an update of the current evaluation if PUSD is able to provide the relevant data points.

⁵ While four out of five cohorts in the analysis enter PUSD in Grade K, we use Grade 1 demographic characteristics which are available for all students in the sample and are measured sufficiently early for any potential effect of IB PYP participation/non-participation to manifest itself.

OUTCOMES OF INTEREST

In this report, Hanover looks at the following middle school outcomes to compare post-program performance between participants and matched similar non-participants. All outcomes are evaluated in Grades 6, 7, and 8.

- **English Skills Acquisition**
 - Reclassification Status⁶
 - Number of Years in PUSD EL Program⁷
- **Academic Outcomes**
 - SBAC ELA Achievement Level (Met Standard and Exceeded Standard)⁸
 - SBAC ELA Scaled Score
 - SBAC ELA Research/Inquiry Claim Achievement Level (Met Standard and Exceeded Standard)
 - SBAC Math Achievement Level (Met Standard and Exceeded Standard)
 - SBAC Math Scaled Score
 - GPA
- **Behavioral Outcomes**
 - Attendance Rate
 - Probability of At Least One Disciplinary Incident
 - Probability of At Least One Out-of-School Suspension

STATISTICAL METHODOLOGY

Hanover uses **regression analysis** to compare outcomes for IB PYP participants and *matched* non-participants who enroll in a non-IB middle school. While Hanover accounts for most effects of student demographics through the PSM procedure, there may be other potential sources of bias, particularly the effects of attending a particular middle school or belonging to a particular student cohort. Regression analysis allows Hanover to isolate the pure effect of the IB PYP while controlling for these observable factors.

Hanover uses **linear regression** to evaluate the effect of the program on continuous outcomes (e.g. GPA). In these models, the coefficients are interpreted as an average difference between outcomes for IB PYP participants and matched non-participants. Hanover also uses **logistic regression (logit)** to estimate the effect of the program on binary outcomes (e.g. probability of meeting standard on SBAC ELA assessment). In this case, for the ease of interpretation, Hanover reports the marginal effect at means (MEMs).

⁶ Reclassification Status for English Learners is to compare all reclassified ever-EL students with the current EL students enrolled in each grade level of analysis.

⁷ Number of years in district EL program is calculated only for students who are reclassified at the grade level of observation (i.e. those students for whom the years in EL have already stopped accumulating).

⁸ Met Standard is a binary variable that has a value of 1 if student Met or Exceeded Standard. Exceeded Standard has a value of one if student Exceeded Standard.

Due to small sample size, regression analysis is not feasible for IB PYP and matched non-participants who enroll in Blair HS during their middle school years. In this part of the analysis, Hanover conducts **t-tests** to compare the differences in outcomes between the two groups and calculate its statistical significance without controlling for the school and cohort effects.

SECTION II: EFFECTS OF IB PYP

In this section, Hanover compares the middle school academic and behavioral performance for IB PYP and *matched similar* non-participants. The analysis is conducted using regression framework and statistical t-tests.

EFFECTS FOR NON-IB MIDDLE SCHOOL STUDENTS

MATCHING PROCEDURE

To construct a comparison group, Hanover employs the PSM technique and match the IB PYP participants **who enrolled in non-IB middle school** to the most similar non-participants based on demographic characteristics (gender, race/ethnicity, Special Education status, economic disadvantage status⁹, English Learner status). Overall, the PSM procedure allows us to match 232 IB PYP participants to 232 similar non-participants.¹⁰

Figure 3.1 presents the results of t-tests on the outcomes of interest for the matched groups, without controlling for students’ school or cohort. **Figures 3.2-3.8** present the results of regression analysis on the outcomes of interest. The results are organized by the type of outcome. The t-tests and the regression analysis generally agree, with the significance tests matching nearly perfectly and the magnitude of the differences between groups also similar.

Figure 3.1: Summary Statistics on Outcome Variables – by Program Participation Status

MIDDLE SCHOOL OUTCOMES	ELIGIBLE NON-PARTICIPANTS		IB PYP PARTICIPANTS		DIFFERENCE
	N	MEAN	N	MEAN	
Grade 6 Outcomes					
SBAC ELA Met Standard	159	34%	161	30%	-4%
SBAC ELA Exceeded Standard	159	12%	161	9%	-3%
SBAC ELA Scaled Score	159	2486.65	161	2492.24	5.59
SBAC ELA Research/Inquiry Claim Met Standard	159	69%	161	73%	3%
SBAC ELA Research/Inquiry Claim Exceeded Standard	159	19%	161	21%	2%
SBAC Math Met Standard	159	24%	157	22%	-2%
SBAC Math Exceeded Standard	159	13%	157	12%	-1%
SBAC Math Scaled Score	159	2473.66	157	2479.36	5.7
GPA	167	2.67	164	2.59	-0.08
Attendance Rate	167	97%	165	98%	1%***

⁹ When Hanover conducts PSM for students who enrolled in non-IB middle schools, economic disadvantage status is not included in the PSM procedure for Cohort 1 students because of irregularity in this variable in 2008-09 school year. When Hanover conducts PSM for students who enrolled in IB middle school (Blair HS), Hanover does not include economic disadvantage status in any of the cohorts of students.

¹⁰ Although there are 238 participants who enrolled in non-IB middle school in the final analytic dataset, only 232 of them get matched to a similar non-participant after the PSM. The potential reasons are 1) students with missing values in certain variables used in PSM are excluded; and 2) participants with could not be matched to a non-participant with a similar “propensity score” are excluded.

MIDDLE SCHOOL OUTCOMES	ELIGIBLE NON-PARTICIPANTS		IB PYP PARTICIPANTS		DIFFERENCE
	N	MEAN	N	MEAN	
Prob. of Incident	172	6%	178	8%	3%
Prob. of OSS	172	5%	178	8%	3%
Reclassification Rate	96	74%	99	79%	5%
Years in EL	71	3.37	78	3.23	-0.14
Grade 7 Outcomes					
SBAC ELA Met Standard	172	30%	174	37%	0.07
SBAC ELA Exceeded Standard	172	6%	174	13%	0.06**
SBAC ELA Scaled Score	172	2498.23	174	2521.43	23.20**
SBAC ELA Research/Inquiry Claim Met Standard	171	60%	174	63%	3%
SBAC ELA Research/Inquiry Claim Exceeded Standard	171	13%	174	22%	10%**
SBAC Math Met Standard	171	23%	175	23%	0%
SBAC Math Exceeded Standard	171	11%	175	13%	2%
SBAC Math Scaled Score	171	2480.15	175	2487.41	7.25
GPA	177	2.58	179	2.58	0
Attendance Rate	177	97%	181	98%	1%**
Prob. of Incident	178	10%	183	5%	-4%
Prob. of OSS	178	7%	183	5%	-2%
Reclassification Rate	94	76%	92	79%	4%
Years in EL	71	3.41	73	3.19	-0.22
Grade 8 Outcomes					
SBAC ELA Met Standard	130	33%	129	40%	7%
SBAC ELA Exceeded Standard	130	5%	129	12%	6%*
SBAC ELA Scaled Score	130	2516.75	129	2543.7	26.94**
SBAC ELA Research/Inquiry Claim Met Standard	130	63%	128	75%	12%**
SBAC ELA Research/Inquiry Claim Exceeded Standard	130	15%	128	21%	6%
SBAC Math Met Standard	127	26%	129	22%	-4%
SBAC Math Exceeded Standard	127	9%	129	13%	4%
SBAC Math Scaled Score	127	2496.96	129	2511	14.04
GPA	127	2.52	126	2.48	-0.04
Attendance Rate	132	97%	133	97%	0%
Prob. of Incident	134	10%	135	6%	-4%
Prob. of OSS	134	9%	135	6%	-3%
Reclassification Rate	68	82%	68	96%	13%**
Years in EL	56	3.46	65	3.54	0.07

ENGLISH SKILLS ACQUISITION

As seen in **Figure 3.2**, **English Learners who participated in IB PYP are significantly more likely to be reclassified by Grade 8 as compared to non-participants**. The reclassification rate among the former is 16 percentage points higher. At the same time, Hanover does not detect any significant differences between the two groups in Grade 6 and 7 both in terms of reclassification rates and average years that reclassified English Learners spend in the district EL program.

Figure 3.2: Regression Results on Post-Program Student Performance – English Learners

VARIABLE	GRADE 6		GRADE 7		GRADE 8	
	RECLASSIFICATION RATE	YEARS IN EL	RECLASSIFICATION RATE	YEARS IN EL	RECLASSIFICATION RATE	YEARS IN EL
Program Participation Status (Ref: Matched Non-Participants)						
Participants	0.0420	-0.1218	0.0970	-0.2349	0.1643***	0.1647
Observations	195	149	186	144	136	121
R²	--	0.2624	--	0.1285	--	0.0862

Note: * p < 0.1; ** p < 0.05; *** p < 0.01. Additional controls include school and cohort fixed effects. The sample size can vary by model depending on availability of each data point.

ACADEMIC PERFORMANCE

Figures 3.3-3.6 present the effects of IB PYP program on academic outcomes. As a general pattern, **IB PYP has a positive effect on ELA outcomes, which manifests most prominently in Grade 7**.

IB PYP participants are found to be seven to nine percentage points more likely to exceed standards on the SBAC ELA assessment and 28-30 points higher average scaled scores. At the same time, they significantly outperform matched non-participants in Research/Inquiry Claim of the ELA assessment. Specifically, IB PYP participants are 15 percentage points more likely to exceed the standard in Grade 7 and 13 percentage points more likely to meet the standard in Grade 8.

Figure 3.3: Regression Results on Post-Program Student Performance – SBAC ELA

VARIABLE	GRADE 6			GRADE 7			GRADE 8		
	MET	EXCEEDED	SCALED SCORE	MET	EXCEEDED	SCALED SCORE	MET	EXCEEDED	SCALED SCORE
Program Participation Status (Ref: Matched Non-Participants)									
Participants	-0.0141	-0.0163	8.7804	0.0844	0.0911**	30.2229**	0.0710	0.0716*	28.8263**
Observations	320	320	320	346	346	346	259	259	259
R²	--	--	0.0089	--	--	0.0432	--	--	0.06

Note: * p < 0.1; ** p < 0.05; *** p < 0.01. Additional controls include school and cohort fixed effects. The sample size can vary by model depending on availability of each data point.

Figure 3.4: Regression Results on Post-Program Student Performance – SBAC ELA Research/Inquiry Claim Achievement Levels

VARIABLE	GRADE 6		GRADE 7		GRADE 8	
	MET	EXCEEDED	MET	EXCEEDED	MET	EXCEEDED
Program Participation Status (Ref: Matched Non-Participants)						
Participants	0.0159	0.0007	0.0357	0.1489***	0.1280**	0.0622
Observations	320	320	345	345	258	258
R ²	--	--	--	--	--	--

Note: * p < 0.1; ** p < 0.05; *** p < 0.01. Additional controls include school and cohort fixed effects. The sample size can vary by model depending on availability of each data point.

Hanover does not, however, identify any significant differences between participants’ and non-participants’ outcomes in SBAC Math or in their middle school GPAs, which suggests that the primary effect of the program is concentrated in ELA.

Figure 3.5: Regression Results on Post-Program Student Performance – SBAC Math

VARIABLE	GRADE 6			GRADE 7			GRADE 8		
	MET	EXCEEDED	SCALED SCORE	MET	EXCEEDED	SCALED SCORE	MET	EXCEEDED	SCALED SCORE
Program Participation Status (Ref: Matched Non-Participants)									
Participants	0.013	0.0145	11.5547	0.0476	0.0477	17.3075	-0.0575	0.0416	13.8249
Observations	316	316	316	346	346	346	256	256	256
R ²	--	--	0.0156	--	--	0.0176	--	--	0.0264

Note: * p < 0.1; ** p < 0.05; *** p < 0.01. Additional controls include school and cohort fixed effects. The sample size can vary by model depending on availability of each data point.

Figure 3.6: Regression Results on Post-Program Student Performance – GPA

VARIABLE	GRADE 6 GPA	GRADE 7 GPA	GRADE 8 GPA
Program Participation Status (Ref: Matched Non-Participants)			
Participants	-0.1936	0.0052	-0.0056
Observations	331	356	253
R ²	0.0243	0.0141	0.0119

Note: * p < 0.1; ** p < 0.05; *** p < 0.01. Additional controls include school and cohort fixed effects. The sample size can vary by model depending on availability of each data point.

BEHAVIORAL PERFORMANCE

Figure 3.7 presents the effects of the IB PYP program on behavioral outcomes. **The program has a small positive effect on student attendance in Grade 6 and 7.** The average attendance rate for participants is 0.6-0.9 percentage points higher than for non-participants. At the same time, there are no significant differences in the probabilities of disciplinary incidents or out-of-school suspension for students in both groups.

Figure 3.7: Regression Results on Post-Program Student Performance – Behavioral Outcomes

VARIABLE	GRADE 6			GRADE 7			GRADE 8		
	ATT. RATE	INCIDENT	OSS	ATT. RATE	INCIDENT	OSS	ATT. RATE	INCIDENT	OSS
Program Participation Status (Ref: Matched Non-Participants)									
Participants	0.0088*	0.0246	0.0246	0.0063*	-0.0387	-0.0190	-0.0077	-0.0207	-0.0165
Observations	332	350	350	358	361	361	265	269	269
R²	0.0263	--	--	0.0378	--	--	0.0215	--	--

EFFECTS FOR BLAIR HS MIDDLE SCHOOL STUDENTS

In a separate analysis, Hanover performs a similar PSM procedure for IB PYP participants and non-participants **who attend Blair HS in their middle school years.** Through matching on the same set of pre-program characteristics, Hanover pairs 59 IB PYP participants with 59 similar non-participants.

Figure 3.8 presents the results of t-tests on the outcomes of interest for the matched groups. In contrast to the previous analysis, IB PYP participants who continue in the IB program at Blair HS have a slightly lower attendance rate and higher probability of having at least one disciplinary incident or OSS in Grade 6 but still significantly outperform similar non-participants in SBAC ELA outcomes in both Grade 7 and Grade 8, as well as outperform their similar non-participating peers in SBAC Math outcomes in Grade 8. For example, in Grade 7, the likelihood of meeting the SBAC ELA standard for them is 31 percentage points higher, while their probability of exceeding the standard on the SBAC ELA Research/Inquiry Claim is 18 percentage points higher.

The results presented in this subsection must be interpreted with caution due to the small available sample size. Hanover recommends that in a few years PUSD runs a more comprehensive evaluation of outcomes for IB PYP students who attend Blair HS in middle school as the available sample is likely to increase by that time. As shown in **Figure 3.8**, a very small share of students in the cohorts included in this analysis enrolled in Blair HS after completing the IB PYP at Willard ES. However, in the later cohorts, the IB MYP seemingly became a more popular option for new middle school students.

Figure 3.8: Summary Statistics on Outcome Variables – by Program Participation Status¹¹

MIDDLE SCHOOL OUTCOMES	ELIGIBLE NON-PARTICIPANTS		IB PYP PARTICIPANTS		DIFFERENCE
	N	MEAN	N	MEAN	
Grade 6 Outcomes					
SBAC ELA Met Standard	52	56%	51	71%	15%
SBAC ELA Exceeded Standard	52	17%	51	24%	6%
SBAC ELA Scaled Score	52	2529.31	51	2552.88	23.57
SBAC ELA Research/Inquiry Claim Met Standard	51	84%	51	88%	4%
SBAC ELA Research/Inquiry Claim Exceeded Standard	51	29%	51	39%	10%
SBAC Math Met Standard	52	42%	51	57%	15%
SBAC Math Exceeded Standard	52	23%	51	16%	-7%
SBAC Math Scaled Score	52	2513.54	51	2537.88	24.34
GPA	52	3.12	51	3.12	0.01
Attendance Rate	52	98%	51	97%	-1%**
Prob. of Incident	52	0%	51	8%	8%**
Prob. of OSS	52	0%	51	8%	8%**
Reclassification Rate	9	--	9	--	--
Years in EL	8	--	6	--	--
Grade 7 Outcomes					
SBAC ELA Met Standard	43	47%	36	78%	31%***
SBAC ELA Exceeded Standard	43	14%	36	22%	8%
SBAC ELA Scaled Score	43	2544.05	36	2590.28	46.23**
SBAC ELA Research/Inquiry Claim Met Standard	43	77%	36	86%	9%
SBAC ELA Research/Inquiry Claim Exceeded Standard	43	23%	36	42%	18%*
SBAC Math Met Standard	43	40%	36	53%	13%
SBAC Math Exceeded Standard	43	16%	36	19%	3%
SBAC Math Scaled Score	43	2519.6	36	2549.69	30.09
GPA	43	2.86	36	3.01	0.16
Attendance Rate	43	98%	36	97%	-1%
Prob. of Incident	43	2%	36	3%	0%
Prob. of OSS	43	2%	36	3%	0%
Reclassification Rate	11	91%	5	--	--
Years in EL	10	--	4	--	--
Grade 8 Outcomes					
SBAC ELA Met Standard	33	52%	26	77%	25%**
SBAC ELA Exceeded Standard	33	15%	26	19%	4%
SBAC ELA Scaled Score	33	2552.48	26	2597.15	44.67*

¹¹ Note that t-tests are not run for outcomes with a sample size of 10 or fewer observation in a given group.

MIDDLE SCHOOL OUTCOMES	ELIGIBLE NON-PARTICIPANTS		IB PYP PARTICIPANTS		DIFFERENCE
	N	MEAN	N	MEAN	
SBAC ELA Research/Inquiry Claim Met Standard	33	70%	25	92%	22%**
SBAC ELA Research/Inquiry Claim Exceeded Standard	33	30%	25	28%	-2%
SBAC Math Met Standard	33	33%	26	50%	17%
SBAC Math Exceeded Standard	33	21%	26	31%	10%
SBAC Math Scaled Score	33	2539.33	26	2589.46	50.13*
GPA	29	2.66	25	2.92	0.26
Attendance Rate	33	98%	26	98%	0%
Prob. of Incident	33	3%	26	0%	-3%
Prob. of OSS	33	3%	26	0%	-3%
Reclassification Rate	7	--	4	--	--
Years in EL	6	--	3	--	--

Note: * p < 0.1; ** p < 0.05; *** p < 0.01. The statistically significant differences are highlighted in red.

PROJECT EVALUATION FORM

Hanover Research is committed to providing a work product that meets or exceeds client expectations. In keeping with that goal, we would like to hear your opinions regarding our reports. Feedback is critically important and serves as the strongest mechanism by which we tailor our research to your organization. When you have had a chance to evaluate this report, please take a moment to fill out the following questionnaire.

<http://www.hanoverresearch.com/evaluation/index.php>

CAVEAT

The publisher and authors have used their best efforts in preparing this brief. The publisher and authors make no representations or warranties with respect to the accuracy or completeness of the contents of this brief and specifically disclaim any implied warranties of fitness for a particular purpose. There are no warranties that extend beyond the descriptions contained in this paragraph. No warranty may be created or extended by representatives of Hanover Research or its marketing materials. The accuracy and completeness of the information provided herein and the opinions stated herein are not guaranteed or warranted to produce any particular results, and the advice and strategies contained herein may not be suitable for every client. Neither the publisher nor the authors shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages. Moreover, Hanover Research is not engaged in rendering legal, accounting, or other professional services. Clients requiring such services are advised to consult an appropriate professional.



4401 Wilson Boulevard, Suite 400

Arlington, VA 22203

P 202.559.0500 F 866.808.6585

www.hanoverresearch.com