



# VANGUARD PROGRAM CAPSTONE REPORT

Prepared for Pasadena Unified School District

April 2020



---

# TABLE OF CONTENTS

INTRODUCTION.....	3
RECOMMENDATIONS.....	3
KEY FINDINGS.....	3
SECTION I: FORMATIVE EVALUATION.....	5
Research Questions .....	5
Methodology.....	5
Findings.....	6
SECTION II: SUMMATIVE EVALUATION.....	10
Research Question .....	10
Methodology.....	10
Findings.....	12

## INTRODUCTION

Pasadena Unified School District (PUSD) requested Hanover Research's support in evaluating its Vanguard program. The program uses the Summit Learning software platform to provide self-guided learning in English language arts (ELA) and mathematics (Math).

PUSD piloted the Vanguard program in 2016-2017 with a limited, select group of students at McKinley Middle School. In the following two years, PUSD expanded the program but continued to offer the program disproportionately to students identified as gifted. PUSD would like to understand how students who participate in the Vanguard program compare and if they demonstrate growth at the same rate as students who participate in the traditional instructional model. Hanover Research completed two projects to evaluate the Vanguard program:



**Formative Evaluation** – This project examined program participation and usage, as well as formative assessment outcomes measured through the Summit Learning platform.



**Summative Evaluation** – This project examined summative program outcomes by using propensity score matching (PSM) to compare outcomes for nonparticipants who were similar to program students, as observed in the data.

## RECOMMENDATIONS

Based on our findings, Hanover Research recommends that PUSD consider the following recommendations.



**Continue to monitor the progress of students who participate in the Vanguard Program.** While propensity score matching results indicate that program participants performed better than nonparticipants in the 2016-2017 and 2018-2019 school years, this was not the case in the 2017-2018 school year.



**Conduct additional research to determine whether particular teachers have more success in implementing the Vanguard program than others.** The analysis would seek to identify practices that teachers are implementing that may be leading to improved outcomes for their students.



**Increase support for non-gifted students in the Vanguard program.** Although performance for gifted Vanguard program participants remained relatively stable from 2016-2017 to 2018-2019, performance for non-gifted participants declined over this period. This may indicate a need to better support non-gifted students in the program.

## KEY FINDINGS



**Overall, SBAC scores are higher for Vanguard participants than other students in PUSD.** When using a propensity score matching (PSM) procedure to compare scores for Vanguard participants to similar students across PUSD, scores are higher for Vanguard students across most years and grade levels. Likewise, SBAC scores for both gifted and non-gifted Vanguard participants at McKinley School were higher than scores for gifted and non-gifted nonparticipants at McKinley School, from the 2016-2017 academic year to the 2018-2019 academic year.



**However, the effect of Vanguard participation on SBAC scores appears to vary across years.** Although scores for Vanguard students were higher across subjects and grade levels than for other students identified using PSM during the 2016-2017 and 2018-2019 school years, scores for Vanguard students during the 2017-2018 school year were lower than scores for students in the control group.



**SBAC scores for Vanguard participants overall declined from the 2016-2017 school year to the 2018-2019 school year, which may reflect the shift in program students from the more engaged and higher-performing students in the pilot year to broader, more representative groups in the two later years.** Performance for gifted Vanguard students remained relatively stable, while performance for non-gifted students declined by:

- ➔ 9.3 scale points in ELA,
- ➔ 6.9 scale points in Math,
- ➔ 9.2 scale points in History, and
- ➔ 3.1 scale points in Science



**The Vanguard program does not appear to have significantly affected behavioral outcomes.** Absenteeism, suspension, and disciplinary incident rates were lower for program participants than for nonparticipants during the program's pilot year in 2016-2017. However, the gaps between program participants and nonparticipants declined between 2016-2017 and 2018-2019 as the Vanguard program expanded to a broader range of students.

# SECTION I: FORMATIVE EVALUATION

This section discusses findings from the formative evaluation. The section begins by outlining the research questions and methodology used to address them before discussing evaluation findings related to Vanguard program enrollment and outcomes.

## RESEARCH QUESTIONS

The formative evaluation addressed the following research questions:



**How does Vanguard program participation and usage differ over time and across subgroups?**



**Are there any changes in the trend of student outcomes over time? Do students in the program perform better on these outcomes than students not in the program?**



**Do students in the program perform better on these outcomes than students not in the program?**

## METHODOLOGY

Figure 1.1 presents the data provided by PUSD to support the formative program evaluation. Hanover Research analyzed data for participants and nonparticipants at McKinley School to examine trends in program enrollment and compare outcomes for Vanguard participants to nonparticipants in the same school.

**Figure 1.1: Formative Evaluation Data Overview**

DATA PROVIDED	INDICATORS
<b>Data Provided for PUSD</b>	
Enrollment	<ul style="list-style-type: none"> <li>▪ Grade Level</li> <li>▪ School Membership</li> </ul>
Demographics	<ul style="list-style-type: none"> <li>▪ Gender</li> <li>▪ Race/Ethnicity</li> <li>▪ Special Education (SPED) Status</li> </ul>
Academic Performance	<ul style="list-style-type: none"> <li>▪ SBAC Math Proficiency Level and Scaled Scores</li> <li>▪ SBAC ELA Proficiency Level and Scaled Scores</li> </ul>
Attendance and Behavioral Outcomes	<ul style="list-style-type: none"> <li>▪ Days Enrolled</li> <li>▪ Days Present</li> <li>▪ Number of Disciplinary Incidents</li> <li>▪ Number of Suspensions</li> </ul>
<b>Data Provided for McKinley Middle School</b>	
Program Participation and Program Outcomes	<ul style="list-style-type: none"> <li>▪ Overall Score</li> <li>▪ Completion Rate</li> <li>▪ Project Count</li> </ul>

## FINDINGS

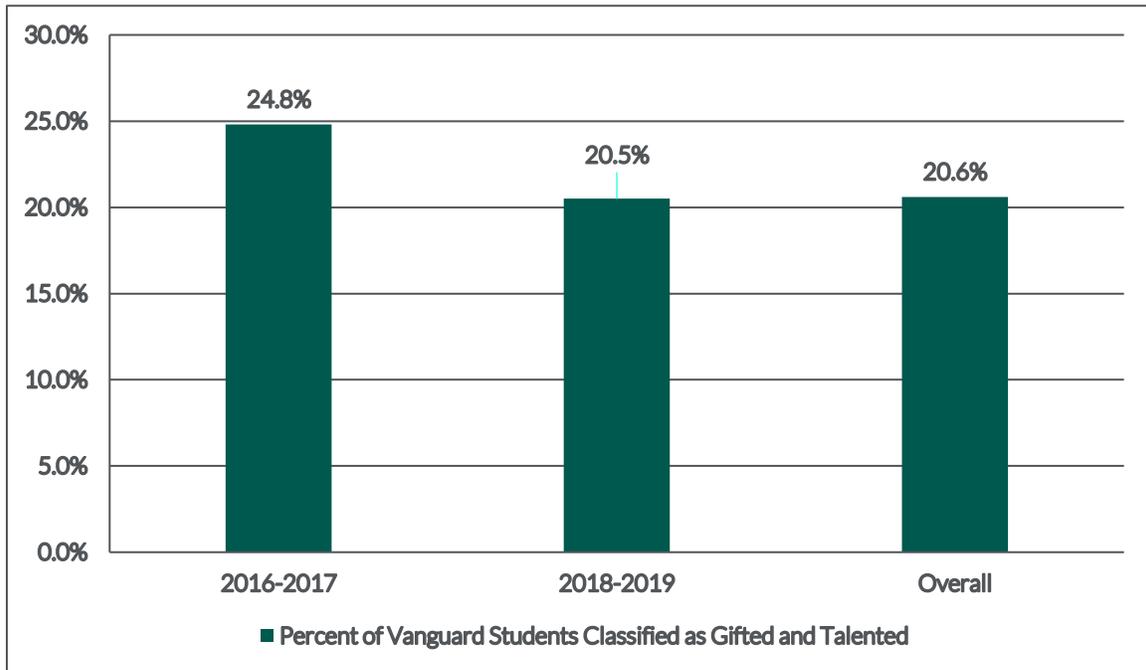
Below, findings are presented related to program enrollment; performance outcomes within the Vanguard program; and outcomes related to attendance, discipline, and Smarter Balanced Assessment (SBAC) scores.

### PROGRAM ENROLLMENT



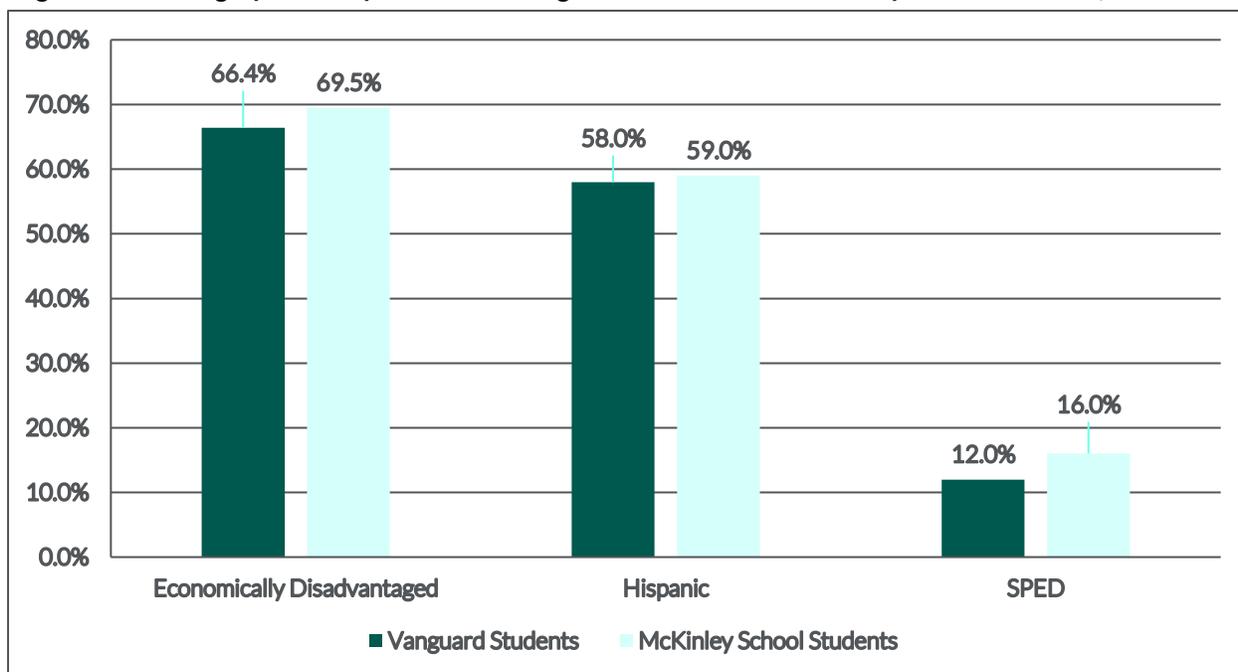
Over time, the profile of Vanguard program students is becoming increasingly reflective of McKinley School and district averages at PUSD. As shown in Figure 1.2, the percent of students classified as gifted and talented (GT) declined from the pilot year of the program in 2016-2017 to 2017-2018.

Figure 1.2: Percent of Vanguard Students Classified as Gifted and Talented



Likewise, the proportion of Vanguard students classified as eligible for special education (SPED), English learners (EL), and economically disadvantaged has increased over time. As shown in Figure 1.3 on the following page, the percentages of Vanguard students classified as economically disadvantaged, Hispanic, and SPED services were similar to percentages for McKinley School by the 2018-2019 school year. However, the gender gap within the Vanguard program is 11.2 percentage points in favor of male students, compared to 4.6 percentage points for McKinley School. The proportion of male students in the Vanguard program increased by 2.6 percentage points from 2016-2017 to 2018-2019.

Figure 1.3: Demographic Comparisons for Vanguard Students and McKinley School Students, 2018-2019



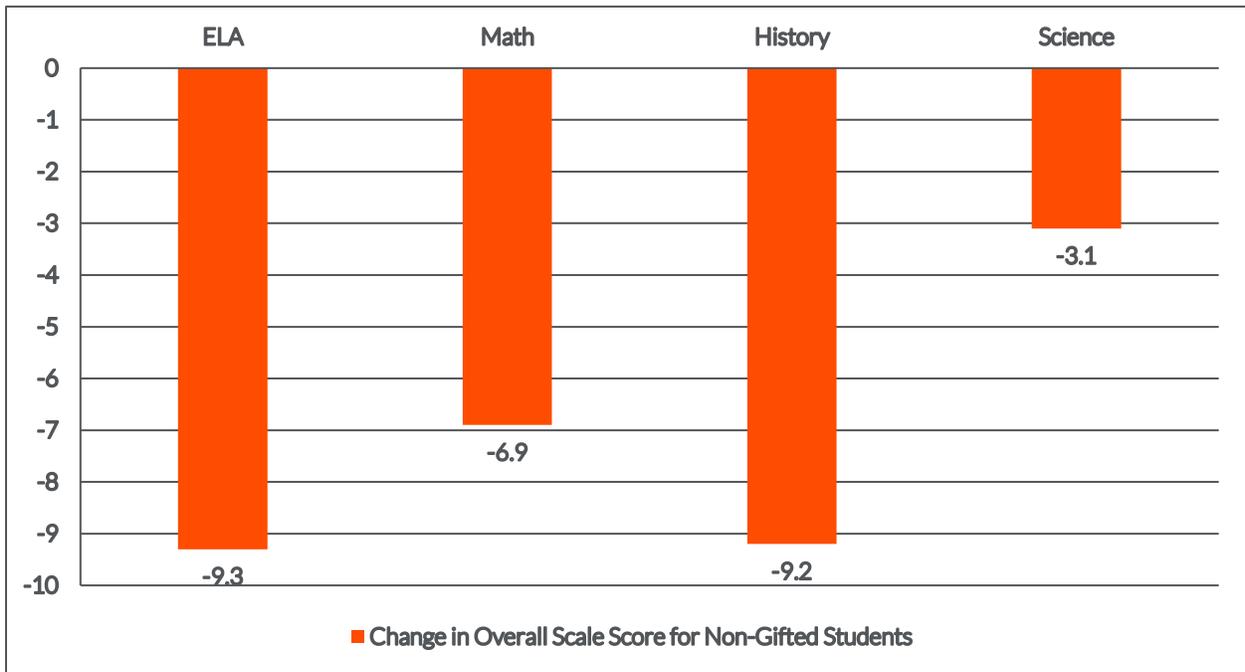
### PROGRAM OUTCOMES



Performance within the program has declined over time. This change may reflect the shift in program students from the more engaged and higher-performing students in the pilot year to broader, more representative groups in the two later years. For example, performance has remained mostly stable among gifted students. The overall math score for gifted students decreased by 1.2 points from 2017-2017 (87.4) to 2018-2019. Overall, ELA, history, and science scores increased slightly for gifted students over the same period.

In contrast, overall scores for non-gifted students declined across all subjects from 2016-2017 to 2018-2019, as shown in Figure 1.4 on the following page. This may indicate a need to better support non-gifted students in the program. Deeper analyses are needed to identify whether the decline is due to the expansion of the program, the type of non-gifted students entering the program, or a lack of resources for non-gifted students in the program.

Figure 1.4: Change in Vanguard Scale Scores for Non-Gifted Participants, 2016-2017 – 2018-2019



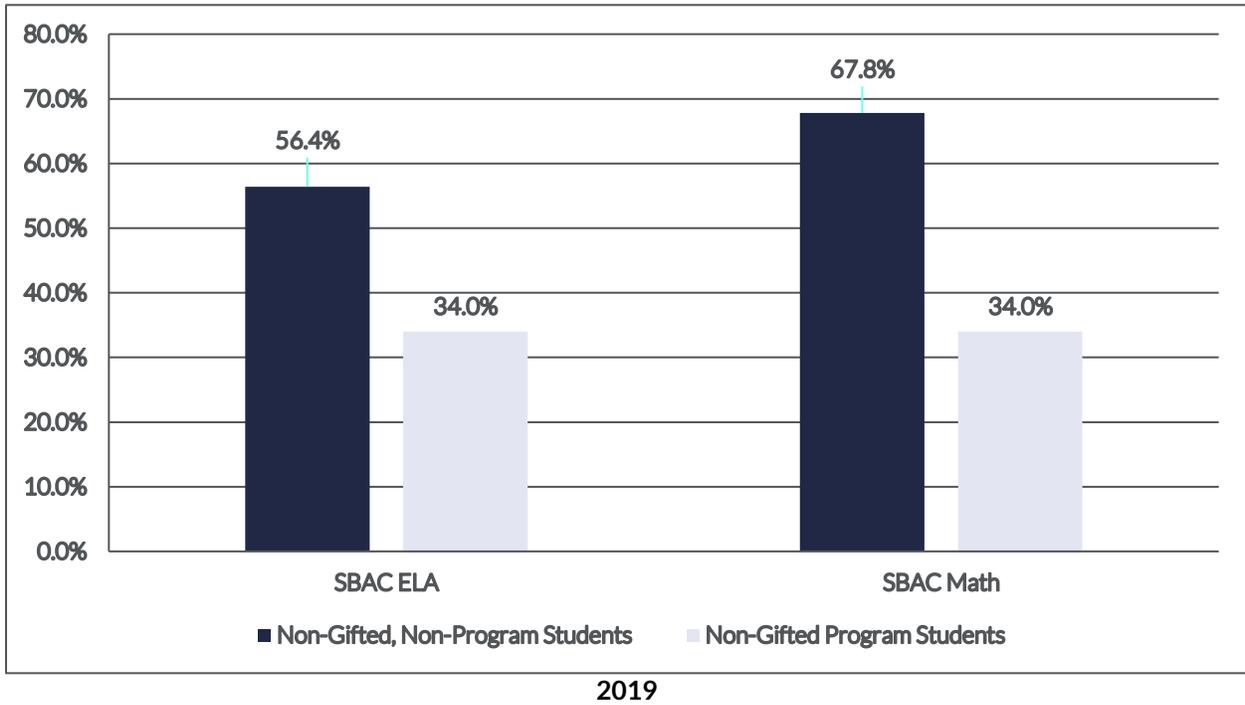
### SBAC, ATTENDANCE, DISCIPLINE



Both gifted and non-gifted students in the Vanguard program scored higher on the SBAC ELA and math assessments than other McKinley students in the same grade. For example, the average SBAC scale score is 2467 among non-gifted non-program McKinley students, 37 points lower than the average score among non-gifted McKinley students (2504). At the same time, the average SBAC math scaled score is 2444 among non-gifted, non-program McKinley students and 44 points higher (2488) among non-gifted program students.

In addition to higher scale scores, the proportion of program students who placed in Level 1 on the SBACs were substantially lower than the proportion of non-program McKinley students, as shown in Figure 1.5 on the following page. This result indicates that the program might be especially helpful for lower-performing students to improve their SBAC outcomes.

Figure 1.5: Percent of Non-Gifted Program and Non-Program Students Placing in SBAC Level 1, 2018-



However, the data provided by PUSD for the formative evaluation does not support causal inferences regarding the Vanguard program's effects on SBAC scores. Non-gifted students enrolled in the Vanguard program may have already been achieving at higher levels than non-program students. More rigorous comparisons across similar groups of students may help infer a causal relationship between program participation and summative assessment outcomes.

As the Vanguard program expanded to a broader group of students at McKinley School from 2016-2017 to 2018-2019, the gaps in absence, incident, and suspension rates between program and non-program students diminished. For example, 2016-2017 students had an average absence rate of 2.9 percent, compared to 6.7 percent among non-program students. In 2018-2019, the gap in absence rates was smaller between program (4.0%) and non-program students (6.1%). There are similar trends in students' rates of disciplinary incidents and suspensions.

## SECTION II: SUMMATIVE EVALUATION

This section presents findings from the summative program evaluation.

### RESEARCH QUESTION

The summative evaluation examined the following research question:



**How do the SBAC scale scores of Vanguard program participants compare to those of similar nonparticipants in PUSD?**

### METHODOLOGY

The summative evaluation builds off the descriptive analysis by investigating the causal relationship between program participation and SBAC outcomes using a propensity score analysis.

#### PROPENSITY SCORE MATCHING



This analysis sought to identify the impact of the Vanguard program on SBAC performance by estimating how program participants *would* have performed if they had *not* participated in the program and comparing that to their actual performance. In order to do this, Hanover attempted to identify a proxy group of similar non-program participants to serve as a comparison for the program participants.

A simple comparison of participants to all nonparticipants could result in a comparison group inherently dissimilar to participants in ways that may be related to student performance. Therefore, a more appropriate approach is to identify a group of non-participating students who are similar to participants in terms of characteristics that may be related to both likelihood of program participation and academic performance. Such a group would serve as a more direct proxy for the outcomes that the program participants may have experienced had they not been in the program.

To identify a comparison group of nonparticipants, Hanover used a method called propensity score matching (PSM). In this approach, Hanover calculates the likelihood (termed a propensity score) that each given student would participate in the program using a logistic regression framework with a set of predictors described below. Hanover then matches nonparticipants to participants based on their propensity scores, and only the nonparticipant with the closest propensity score to a participant's is matched. Therefore, the constructed comparison group comprises nonparticipants who appear to be about as likely to have been in the program as actual program participants. This approach identifies a different matched group for each subject and grade combination.

Propensity score matching procedures operate under the assumption that, after accounting for differences in the matching model variables, program participants are not dissimilar from nonparticipants in ways that may be meaningfully related to student performance. **It should be noted that Hanover's model cannot control for unobserved characteristics and factors that may have an impact on student performance.**

#### MATCHING VARIABLES



To estimate program treatment effects, Hanover identified sets of students who were similar to program participants using PSM. The PSM models were run for each combination of grade and year the program has been available. Hanover ran separate models for the SBAC ELA and SBAC Math scale score outcomes. Grade 6 and Grade 7 students have models for the 2016-17, 2017-18, and 2018-19 academic years; Grade 8 students have models for the 2017-18 and

2018-19 academic years. Hanover used the following variables in the PSM model to determine a student's probability of participating in the program.

- SBAC Scale Score<sup>1</sup>
- Gender
- Gifted Status
- SPED Status
- Economically Disadvantaged<sup>2</sup>
- Suspension in Previous Year<sup>3</sup>

Comparability of the program students and the matched group was assessed using independent sample t-tests for the SBAC scale scores and two sample tests of proportions for each of the dichotomous demographic variables.

## VANGUARD REPEATERS



**Figure 2.1** reports the number of years that Vanguard program participants had previously participated in the Vanguard program. The matching design is intended to match program participants to nonparticipants. Therefore, the matching models did not include years of participation as a matching variable in the PSM. As described in the previous section, the matching models did include students' previous year's SBAC score as a matching variable. To the extent that previous Vanguard participation affected student performance, this effect from prior years should be accounted for in the previous year's SBAC score. The interpretation of treatment effects from this analysis, therefore, reflects the effect of the program associated with one year's worth of participation.

**Figure 2.1: Distribution of Vanguard Repeaters**

YEAR	YEARS IN PROGRAM	GRADE 6	GRADE 7	GRADE 8
2016-17	First Year	60	57	NA
	<b>Total</b>	<b>60</b>	<b>57</b>	<b>NA</b>
2017-18	First Year	75	14	6
	Second Year	NA	54	48
	<b>Total</b>	<b>75</b>	<b>68</b>	<b>54</b>
2018-19	First Year	102	36	10
	Second Year	NA	57	9
	Third Year	NA	NA	45
	<b>Total</b>	<b>102</b>	<b>93</b>	<b>64</b>

## TREATMENT EFFECT MODEL ESTIMATION

Following the matching procedure, we estimated regression models to calculate the mean difference between the treatment group and the matched comparison group. These models controlled for the SBAC scale score in the previous academic year and whether the student was Hispanic<sup>4</sup>. Hanover calculated treatment effect estimation models for each combination of year and grade level.

<sup>1</sup> There were separate models for ELA and math SBAC scores. Eight grade year combinations, and sixteen models total.

<sup>2</sup> Economically disadvantaged was only included in the grade 6 and 7 models. The decision to exclude it in the grade 8 models was due to data limitations that limited the number of covariates in the PSM models.

<sup>3</sup> Suspension in previous year is a dichotomous variable that indicates whether a student was ever suspended in the previous academic year. This indicator replaced the economically disadvantaged indicator as it generated a better match on the SBAC scale score.

<sup>4</sup> Hanover included the Hispanic indicator in the analytic model as sample size constraints prohibited the use of the Hispanic variable in the matching model. Hanover included prior SBAC performance in both the matching and analytic model to minimize the bias of program selection effects that may have been related to student performance.

## EFFECT SIZE CALCULATION



To assess the magnitude of differences in outcomes between Vanguard program participants and nonparticipants, Hanover calculated the effect size (i.e., the standardized difference) for treatment effect model estimates. The purpose of the effect size is to enable direct comparisons of group differences across different outcome types. The effect size also allows for the determination of whether the difference between the two groups is meaningful. Even if a difference is found to be statistically significant, a low effect size would indicate that it is not practically significant. Figure 2.2 presents effect size values and their interpretation as used by Hanover in this analysis.<sup>5</sup> Positive values indicate instances in which program participants outperform nonparticipants, while negative values indicate the opposite.

**Figure 2.2: Effect Size Interpretation**

EFFECT SIZE CATEGORY	ABSOLUTE VALUE OF EFFECT SIZE
Small	0 to less than 0.05
Medium	0.05 to less than 0.20
Large	0.20 or greater

## FINDINGS



In some years and grade levels, Vanguard program participants had higher mean SBAC scale scores than similar nonparticipants. In the 2016-2017 academic year, Grade 6 students who participated in the program outperformed Grade 6 nonparticipants for both the ELA and math assessments. In the 2018-2019 academic year, Grade 7 and 8 students who participated in the program outperformed Grade 7 and 8 nonparticipants for both the ELA and math assessments.

In contrast, Vanguard participants tended to have lower SBAC scale scores than similar nonparticipants during the 2017-2018 school year. In particular, Grade 6 Vanguard participants underperformed Grade 6 nonparticipants on the ELA SBAC, while Vanguard participants in both Grades 6 and 7 underperformed nonparticipants on the math SBAC.

**Figure 2.3: Treatment Effect Estimation, SBAC ELA and Math Scale Scores**

SUBJECT	LARGE POSITIVE EFFECT	SUBSTANTIAL NEGATIVE EFFECT
ELA	Grade 6 (2016-17); Grade 7 (2018-19); Grade 8 (2018-19)	Grade 6 (2017-18)
Math	Grade 6 (2016-17); Grade 7 (2018-19); Grade 8 (2018-19)	Grade 6 (2017-18); Grade 7 (2017-18)

\*Reported effects in this table are statistically significant at a 95% confidence interval or above.

<sup>5</sup> Lipsey, Mark W. et al. "Translating the Statistical Representation of the Effects of Education Interventions into More Readily Interpretable Forms." National Center for Special Education Research. Institute of Education Sciences. November 2012.

# ABOUT HANOVER RESEARCH

Hanover Research provides high-quality, custom research and analytics through a cost-effective model that helps clients make informed decisions, identify and seize opportunities, and heighten their effectiveness.

## OUR SOLUTIONS

### ACADEMIC SOLUTIONS

- **College & Career Readiness:**  
Support on-time student graduation and prepare all students for post-secondary education and careers.
- **Program Evaluation:**  
Measure program impact to support informed, evidence-based investments in resources that maximize student outcomes and manage costs.
- **Safe & Supportive Environments:**  
Create an environment that supports the academic, cultural, and social-emotional needs of students, parents, and staff through a comprehensive annual assessment of climate and culture.

### ADMINISTRATIVE SOLUTIONS

- **Family and Community Engagement:**  
Expand and strengthen family and community relationships and identify community partnerships that support student success.
- **Talent Recruitment, Retention & Development:**  
Attract and retain the best staff through an enhanced understanding of the teacher experience and staff professional development needs.
- **Operations Improvement:**  
Proactively address changes in demographics, enrollment levels, and community expectations in your budgeting decisions.

## LEADERSHIP SOLUTION

Build a high-performing administration that is the first choice for students, parents, and staff.

## OUR BENEFITS



### EXPERT

200+ analysts with multiple methodology research expertise



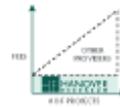
### FLEXIBLE

Ongoing custom research agenda adapts with organizations' needs



### DEDICATED

Exclusive account and research teams ensure strategic partnership



### EFFICIENT

Annual, fixed-fee model shares costs and benefits



[www.hanoverresearch.com](http://www.hanoverresearch.com)