# 2017–2019 Implementation Evaluation of the National Math and Science Initiative's College Readiness Program

NMSI/20160003/Years 2 & 3/Deliverable – March 30, 2021

Julia Phelan, Jeffrey Egger, Junok Kim, Kilchan Choi, Eunhee Keum, Gregory K. W. K. Chung, and Eva L. Baker
CRESST/University of California, Los Angeles

National Center for Research on Evaluation,
Standards, and Student Testing (CRESST)
Center for the Study of Evaluation (CSE)
Graduate School of Education & Information Studies
University of California, Los Angeles
300 Charles E. Young Drive North
GSE&IS Bldg., Box 951522
Los Angeles, CA 90095-1522
(310) 206-1532

The work reported herein was supported by a subgrant from the National Math and Science Initiative, award number 20160003, to the National Center for Research on Evaluation, Standards, and Student Testing (CRESST). The work was made possible by funding from the U.S. Department of Education to the National Math and Science Initiative.

The findings and opinions expressed in this report are those of the authors and do not necessarily reflect the positions or policies of the National Math and Science Initiative or the U.S. Department of Education.

## TABLE OF CONTENTS

Executive Summary ..... 1
I. Introduction ..... 2
II. Program Description ..... 4
III. Research Questions ..... 6
IV. Results for the Impact Study ..... 8
V. Results for Implementation Evaluation ..... 36
VI. Summary ..... 138
VII. References ..... 144
Appendix A: Propensity Score Matching Results ..... 145
Appendix B: Hierarchical Generalized Linear Model ..... 150
Appendix C: 2019 Teacher Survey Training Changes from SY2017–2018 ..... 153
Appendix D: 2019 Teacher Survey Financial Awards Changes from SY2017–2018 ..... 154
Appendix E: 2019 Teacher Survey Preferences for Additional Tools and Materials ..... 155
Appendix F: 2019 Teacher Survey CRP Impact on AP Enrollment ..... 156
Appendix G: 2019 PSD Survey Comments about CRP Matching Expectations ..... 157
Appendix H: 2019 PSD Survey Comments on CRP Impact on School ..... 159
Appendix I: 2019 PSD Survey Factors that Affected CRP Implementation ..... 161
Appendix J: 2019 SC Survey Suggestions for Improving SSS Attendance ..... 162
Appendix K: 2019 SC Survey Factors that Hindered or Assisted CRP Implementation ..... 164

# 2017–2019 IMPLEMENTATION EVALUATION OF THE NATIONAL MATH AND SCIENCE INITIATIVE'S COLLEGE READINESS PROGRAM

Julia Phelan, Jeffrey Egger, Junok Kim, Kilchan Choi, Eunhee Keum, Gregory K. W. K. Chung, and Eva L. Baker
CRESST/University of California, Los Angeles

## Executive Summary

The National Math + Science Initiative (NMSI) is a nonprofit organization committed to improving educational outcomes that traces its roots back to the early 1990s. NMSI's College Readiness Program (CRP) is a long-standing program with the goal of promoting STEM education in high schools to improve students' preparation for college. The three-year program provides teacher, student, and school supports to promote high school students' success in English, mathematics, and science Advanced Placement (AP) courses, with a focus on students who are traditionally underrepresented in the targeted AP courses.

Through a scale-up grant awarded to NMSI by the Investing in Innovation (i3) program, the CRP was implemented in 27 schools in the 2016–2017 school year (Treatment Schools) and in 21 schools in the 2017–2018 school year (Delayed Treatment Schools), collectively identified as Program Schools. CRESST conducted an independent evaluation of the impact of the CRP on students' AP outcomes using a randomized cluster trial with 48 CRP schools and 48 Comparison Schools in 10 states. The evaluation of the CRP consisted of three parts: (1) measuring the program's impact on selected student AP exam outcomes, (2) determining the impact of the program on school perspectives and culture, and (3) assessing of the fidelity of implementation of the CRP at the school level.

AP exam data from 48 Treatment Schools, with a total of 8,778 exams in 2018 and 9,378 in 2019, and 48 matched control schools, with 7,505 exams in 2018 and 2019 in Year 3, were analyzed for this study. Program impact was evaluated using a 2-level hierarchical generalized linear model (HGLM) with students nested within schools. The analyses revealed that in 2018 the probability of a student taking an AP exam in the Program Schools was, on average, 7% higher than the paired Comparison Schools, and the difference was statistically significant. And in 2019 the effect was even greater with the probability of taking an AP exam being significantly higher in Program Schools (18%) than in Comparison Schools (3%). When looking at the probability of an exam yielding a qualifying score, in 2018, the HGLM analyses found no significant difference between the two groups. In 2019, however, exams taken at the Program Schools had a significantly higher overall probability (2%) of receiving a qualifying score than

the Comparison Schools (0%). These analyses compared results to the total school population. In the next analyses, we looked only at those students who took AP exams. In 2018, overall, the fitted probability of achieving a qualifying score among the exams taken was 8% in the Program Schools, compared to 22% in the Comparison Schools. In 2019, however, the difference between the Program Schools (7%) and the Comparison Schools (9%) was not statistically significant.

Given differences in school sizes, program impact was further evaluated using a conditional HGLM. Findings were similar as to those from the unconditional model for the number of students taking AP exams, and for the number of qualifying scores (when looking at the proportion of school population). When only looking at exam results for those who took the exam, the fitted probability of receiving a qualifying score among the exams taken was again higher in the Comparison Schools (as was the case in the unconditional model results). However, the difference between the two groups was not significant with a gap narrowed down from 14% to 10%.

Fidelity of implementation was evaluated using a fidelity matrix approach (required as part of the evaluation of the i3 program), which showed that not all elements of the program were implemented with high fidelity. In 2018, results indicated that 43 out of 48 schools (90%) achieved 80% or better implementation fidelity, for an average fidelity score of 89%. Four schools achieved a perfect 100% fidelity score. In 2019, 88% of schools achived 80% or better implementation fidelity. Ten schools achieved a perfect 100% fidelity score. In 2018 in more than 75% of schools, not all teachers fulfilled their requirements for attending the required teacher training sessions, and so this component was not implemented with fidelity. In 2019, this picture improved a little with 15 schools (31%) meeting the 80% threshold. Teacher stipends, administrator bonuses, and student qualifying score awards were paid as expected.

Teacher survey data indicated that teachers found the training and professional development activities provided by the CRP to be the most benefical program supports. Mentoring was chosen, across all years, as the least effective program component. When prompted for the second most effective CRP component, the same number of teachers selected the funding of classroom and lab supplies as did teacher training.

## I. Introduction

Proficiency in math and science is crucial to our country's capacity for innovation and future economic growth, yet a growing number of students lack foundational knowledge and skills in these subjects. Performance in math and science of U.S. college students is also below that of their peers in many other nations (Chen, 2013; Fleischman et al., 2010). In 2011, for example, roughly one third of U.S. bachelor's degrees were awarded in science and engineering

fields, compared to 60% in Japan and 50% in China (National Science Board, 2014). Indeed, it is estimated that in 2016, only around 41% of U.S. high school graduates were ready for college-level math, and only 36% ready for college-level science (ACT, 2016). The most recent Programme for International Student Assessment (PISA) results, from 2015, found the U.S. placed 38th of 71 countries in math and 34th in science (Pew Research Center, 2017). The necessity for an increased focus on math and science specifically is based on years of research which shows fewer students are entering math- and science-related career fields (National Science Board, 2010).

These issues are even more pronounced for high-need and traditionally underserved students who may face hurdles because of policies and mindsets that limit their ability to access rigorous coursework. Data from the National Science Foundation (NSF) found that 27% of ninth graders in the lowest socioeconomic status category were not enrolled in any science courses, compared with 11% of students in the highest income category. These differences in access and opportunity can lead to achievement gaps that continue through college and beyond. The gap between White students' six-year college graduation rates and their African American peers is around 22%, and the gap between White students and their Hispanic peers is 10% (Kena et al., 2014).

The National Math and Science Initiative (NMSI) was formed to address the declining number of students prepared to take rigorous college courses in math and science and equipped for careers in those fields. The College Readiness Program (CRP) was created to raise the academic bar in public schools by demonstrating that more students, especially high-need students, can master rigorous Advanced Placement (AP®) coursework, with a particular emphasis on math and science. The CRP addresses the need to improve STEM education, increase academic intensity, and improve student achievement in order to decrease the college readiness gap, especially among traditionally underrepresented and high-need students.

Program supports include teacher training and mentoring; additional student instruction time outside the classroom; instructional resources; incentive payments to teachers, administrators, and students (tied to AP performance); and funding for purchasing equipment and supplies. Over a three-year period, the CRP supports existing high schools focused on school reform and changing school culture. The study aimed to explore the impact of NMSI's CRP on selected student outcomes and evaluate the quality of implementation of the program in 15 school districts across 10 states. This report presents findings from the evaluation of the impact of the CRP as well as findings from the fidelity of implementation study based on a series of metrics, the "fidelity matrix."

In the next section we provide an overview of the CRP and key components which were the focus of the analysis of the quality of program implementation in this evaluation.

## II. Program Description

### A. CRP Logic Model

The CRP logic model (Figure 1) presents the key components of the intervention: program management, teacher support, student support, and financial awards. For teachers, the program offers (a) course-specific training, (b) access to expert mentors, (c) curricular resources, and (d) a financial stipend for participating in program activities. For students, the program offers (a) weekend study sessions led by seasoned instructors, (b) exam fee subsidies, and (c) access to classroom and lab materials needed to support rigorous coursework. For schools, the program offers (a) an experienced liason, or Program Manager (PM); (b) performance goals for teachers, students, and schools; and (c) a financial stipend for administering program activities. At all levels there are also financial incentives associated with AP exam performance.

Teacher participation in professional development and mentoring, their access to materials and resources, and the use of incentives are designed to drive increased familiarity with, and use of, pedagogical strategies as well as increased content knowledge and increased effectiveness in the classroom. Those intermediate outcomes should drive longer term outcomes such as a rise in AP course enrollment, more qualifying AP scores, growth in the number of students prepared to continue rigorous courses of study following high school, and an increased number of teachers qualified to teach AP courses.

For students, additional time on task, access to materials, awards for performance, high-quality resources, and exposure to highly trained teachers are all designed to increase student engagement, preparation, and motivation to perform well. These intermediate outcomes should then influence AP enrollment and the number of qualifying AP scores. Furthermore, more students at a school obtaining qualifying scores and having positive experiences in math, science, and English language arts (ELA) AP courses should positively impact the number of students persisting in STEM courses.

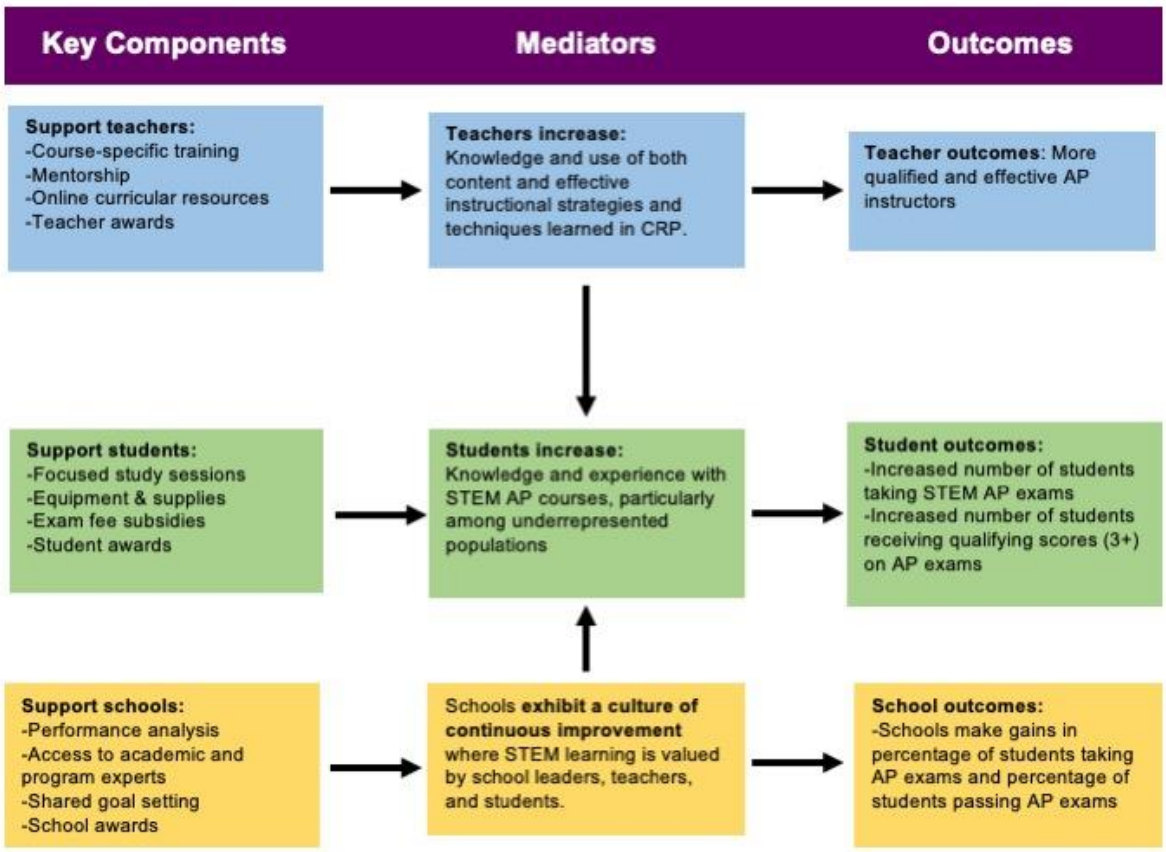



Figure 1. Logic model supporting NMSI's College Readiness Program.

## B. Program Structure

Teacher training begins with an intensive four-day summer session which is reinforced and built upon with three additional days of training throughout the school year. NMSI AP teachers are also given access to an expert mentor to provide coaching and support during the school year including, but not limited to, guidance on pacing, common challenges, and locating additional instructional resources. Most mentor contact is limited to video conferencing and email, but on occasion mentors have visited teachers' classrooms during the academic year to provide opportunities for CRP teachers to observe a more experienced "master teacher." Program teachers gain access to in-depth instructional resources in hard copy through training sessions and student study sessions (SSS), and online from trainers, student study session facilitators, mentors, and a site maintained by NMSI.

To provide more time on task for students, each CRP AP course includes three 4-hour blocks of weekend instruction taught by master AP teachers—time that equates to three extra weeks of AP class time—exposing students to different teaching perspectives and methods. These student study sessions also provide professional development and collaboration

opportunities for teachers who can connect with peers from the region and can see how expert teachers address difficult parts of AP courses. For the three years of NMSI program implementation, teachers continue to receive progressively more rigorous training and lessons; teachers and administrators continue to push further toward increasingly challenging goals; and both students and teachers receive nominal monetary awards for success. NMSI staff work with teachers and administrators throughout implementation to track progress toward their goals and troubleshoot where needed.

## III. Research Questions

The evaluation of the CRP under the i3 scale-up grant consisted of two parts: (1) assessment of the program's impact on selected student AP exam outcomes and (2) assessment of the fidelity of implementation of the CRP.

Research questions for the evaluation were as follows:

1. What was the impact of the CRP on the likelihood that students took STEM-related AP courses?
2. What was the impact of the CRP on the likelihood that students achieved a qualifying score of 3 or higher on STEM-related AP exams?
3. To what extent were each of the key components of the CRP implemented with fidelity?
4. What were the facilitators and barriers to implementation?

In addition to these outcomes, we evaluated the intermediate outcomes and presumptions supporting the logic model. The three-year study was conducted in 48 Program Schools across 11 distinct regions of the United States.

### A. Study Design and Sample

The study began with the staggered introduction of the CRP to 48 schools (Program Schools) in 10 states involved in the i3 scale-up grant. We conducted a randomized cluster trial (RCT), assigning the Program Schools either to treatment or control conditions (Treatment Schools and Delayed Treatment Schools, respectively). The 27 Treatment Schools began implementing the CRP during the first year of the study (the 2016–2017 school year). The Delayed Treatment Schools comprised 21 Program Schools in which implementation of CRP involved a one-year delay and began in the 2017–2018 school year. For 48 matched Comparison Schools, the College Board provided anonymized AP exam data for the second and third years of the study (see Table 1). The Comparison Schools were selected using a propensity score matching technique by the College Board based on the demographic information of the Program Schools. The demographic variables included in the propensity score matching model were

region, the percentage of Black and Hispanic students, the percentage of students eligible for free/reduced price lunch, and number of students enrolled. The results of the propensity score matching provided by the College Board are presented in Appendix A.

Table 1

*Program School and Comparison School Student Populations*

| Program School | Study Group | 2017–2018 Enrollment | Comparison School | 2017–2018 Enrollment |
|---|---|---|---|---|
| California 2 | Treatment School | 1,756 | CompCA 2 | 2,126 |
| California 3 | Delayed Treatment | 1,568 | CompCA 3 | 1,938 |
| Georgia 1 | Treatment School | 722 | CompGA 1 | 1,092 |
| Georgia 2 | Treatment School | 1,205 | CompGA 2 | 1,575 |
| Georgia 3 | Delayed Treatment | 1,332 | CompGA 3 | 1,702 |
| Georgia 4 | Delayed Treatment | 1,889 | CompGA 4 | 2,259 |
| Illinois 1 | Treatment School | 1,014 | CompIL 1 | 1,384 |
| Illinois 2 | Treatment School | 780 | CompIL 2 | 1,150 |
| Illinois 3 | Treatment School | 633 | CompIL 3 | 1,003 |
| Illinois 4 | Delayed Treatment | 652 | CompIL 4 | 1,022 |
| Illinois 5 | Delayed Treatment | 662 | CompIL 5 | 1,032 |
| Louisiana 1 | Treatment School | 1,113 | CompLA 1 | 1,483 |
| Louisiana 2 | Delayed Treatment | 1,025 | CompLA 2 | 1,395 |
| Michigan 1 | Treatment School | 2,458 | CompMI 1 | 2,828 |
| Michigan 2 | Treatment School | 1,177 | CompMI 2 | 1,547 |
| Michigan 3 | Delayed Treatment | 1,985 | CompMI 3 | 2,355 |
| Missouri 1 | Treatment School | 400 | CompMO 1 | 770 |
| Missouri 2 | Treatment School | 602 | CompMO 2 | 972 |
| Missouri 3 | Treatment School | 550 | CompMO 3 | 920 |
| Missouri 4 | Delayed Treatment | 245 | CompMO 4 | 615 |
| Missouri 5 | Delayed Treatment | 583 | CompMO 5 | 953 |
| Missouri 6 | Delayed Treatment | 352 | CompMO 6 | 722 |
| North Dakota 1 | Treatment School | 1,252 | CompND 1 | 1,622 |
| North Dakota 2 | Treatment School | 1,189 | CompND 2 | 1,559 |
| North Dakota 3 | Treatment School | 1,285 | CompND 3 | 1,655 |
| North Dakota 4 | Delayed Treatment | 1,145 | CompND 4 | 1,515 |
| North Dakota 5 | Delayed Treatment | 1,458 | CompND 5 | 1,828 |

| Program School | Study Group | 2017–2018 Enrollment | Comparison School | 2017–2018 Enrollment |
|---|---|---|---|---|
| Ohio 1 | Treatment School | 315 | CompOH 1 | 685 |
| Ohio 2 | Treatment School | 334 | CompOH 2 | 704 |
| Ohio 3 | Treatment School | 999 | CompOH 3 | 1,369 |
| Ohio 4 | Treatment School | 389 | CompOH 4 | 759 |
| Ohio 5 | Delayed Treatment | 415 | CompOH 5 | 785 |
| Ohio 6 | Delayed Treatment | 223 | CompOH 6 | 593 |
| Ohio 7 | Delayed Treatment | 664 | CompOH 7 | 1,034 |
| Ohio 8 | Delayed Treatment | 190 | CompOH 8 | 560 |
| Pennsylvania 1 | Treatment School | 1,226 | CompPA 1 | 1,596 |
| Pennsylvania 2 | Treatment School | 917 | CompPA 2 | 1,287 |
| Pennsylvania 3 | Treatment School | 1,010 | CompPA 3 | 1,380 |
| Pennsylvania 4 | Delayed Treatment | 575 | CompPA 4 | 945 |
| Pennsylvania 5 | Delayed Treatment | 1,167 | CompPA 5 | 1,537 |
| Texas 1 | Treatment School | 487 | CompTX 1 | 857 |
| Texas 2 | Treatment School | 2,795 | CompTX 2 | 3,165 |
| Texas 3 | Treatment School | 1,703 | CompTX 3 | 2,073 |
| Texas 4 | Treatment School | 471 | CompTX 4 | 841 |
| Texas 5 | Treatment School | 1,119 | CompTX 5 | 1,489 |
| Texas 6 | Delayed Treatment | 1,845 | CompTX 6 | 2,215 |
| Texas 7 | Delayed Treatment | 780 | CompTX 7 | 1,150 |
| Texas 8 | Delayed Treatment | 1,778 | CompTX 8 | 2,148 |

## IV. Results for the Impact Study

### A. Descriptive Statistics of 2017–2019 AP Course Taking and Exam Data

We first present the descriptive statistics for the number of AP exams taken as well as AP exam outcomes for the 2017–2018 and 2018–2019 school years. We compared data for science, math, and ELA AP exams between the NMSI CRP schools and the matched comparison sample. We analyzed the number of AP exams taken as well as the number of qualifying scores (score of 3 or higher) achieved at each school.

Data for the Comparison Schools were only available at the aggregate school level and so we compared the number of tests taken and test outcomes at the school level for our comparative analyses.

Note, the data available to us were the number of AP exams taken at each school which is not the same as the *number of students* taking exams at each school; in some cases, students took more than one AP exam, and all exams taken were included in the analyses. Therefore, when we report outcome data, we report the number of exams taken and not the number of students who took them as there is not necessarily a 1:1 correspondence.

School-level AP course enrollment data for the Comparison Schools was not available, so we do not know how many students who took an AP course did not end up taking the AP exam. We do, however, know that in the Program Schools, most students who take AP courses also take the exams. Committing to taking the AP exam is one of the NMSI CRP program requirements which in our experience was followed by most schools and AP teachers. Thus, there are likely more students in the Comparison Schools who take AP courses but do not take the AP exam as is the trend in many schools across the country. In 2018 (for example) only 38% of schools with AP courses had a requirement that students also take the AP exam (Matthews, 2018). If the Comparison Schools did not all require students who took AP courses to take the exam, this may result in a higher percentage of exam takers in the Comparison Schools achieving a qualifying score (QS), particularly if the students who feel more confident they will pass are the ones who take the test.

NMSI's CRP focuses on both increasing the number of students taking AP exams as well as the percentage of those students who achieve a qualifying score. Given the data availability constraints discussed above, we measured participation in the AP program as well as AP exam achievement in several different ways, all of which are included in Figure 2, Figure 3, Figure 4, and Table 2 and are described below.

1. The number and percentage of exams taken at Program and Comparison Schools compared to the overall student population at each
2. The number and percentage of exams taken which resulted in a qualifying score (as compared to the total student population at the school)
3. The percentage of exams resulting in a qualifying score

We first present data for all AP exams (science, ELA, and math) together (see Figure 2). A brief note on data issues and inconsistencies: For purposes of teacher training and student study sessions, NMSI categorizes all computer science AP courses in the math content area. The College Board provided Comparison School computer science exam results in the science content area. Additionally the College Board data were missing information for the 11$^{th}$ grade ELA AP course, English Language and Composition, and all content areas at the school level for which the results were fewer than 10—determined separately for exams and qualifying scores. When evaluating Program Schools against Comparison Schools, the Program School data

reflected the same limiting factors (meaning, for example, that we did not include the 11th grade AP English Language and Composition data so we would have an "apples to apples" comparison). In all other representations of Program School information, the data are inclusive of both low enrollment content area results and English Language and Composition, and computer science is a subject in the math content area.

*Research Question 1: Exams taken as a percentage of the school population:* In both 2018 and 2019 a higher percentage of AP exams were taken in the Program Schools than the Comparison Schools: In 2018, the percentage of AP exams taken in the Program Schools was 6.8% higher than in the comparison group (18.1% vs. 11.3%). In 2019, the percentage of exams taken in the Program Schools increased to 19.4%, while in the Comparison Schools the percentage remained almost the same (11.2%).

*Specific subject areas:* For each of the subject domains (ELA, math, and science) the results showed a similar pattern to total exams. In all subject areas, Program Schools had a higher proportion of exams taken in 2018 and 2019 compared to the Comparison Schools (see Figure 2).

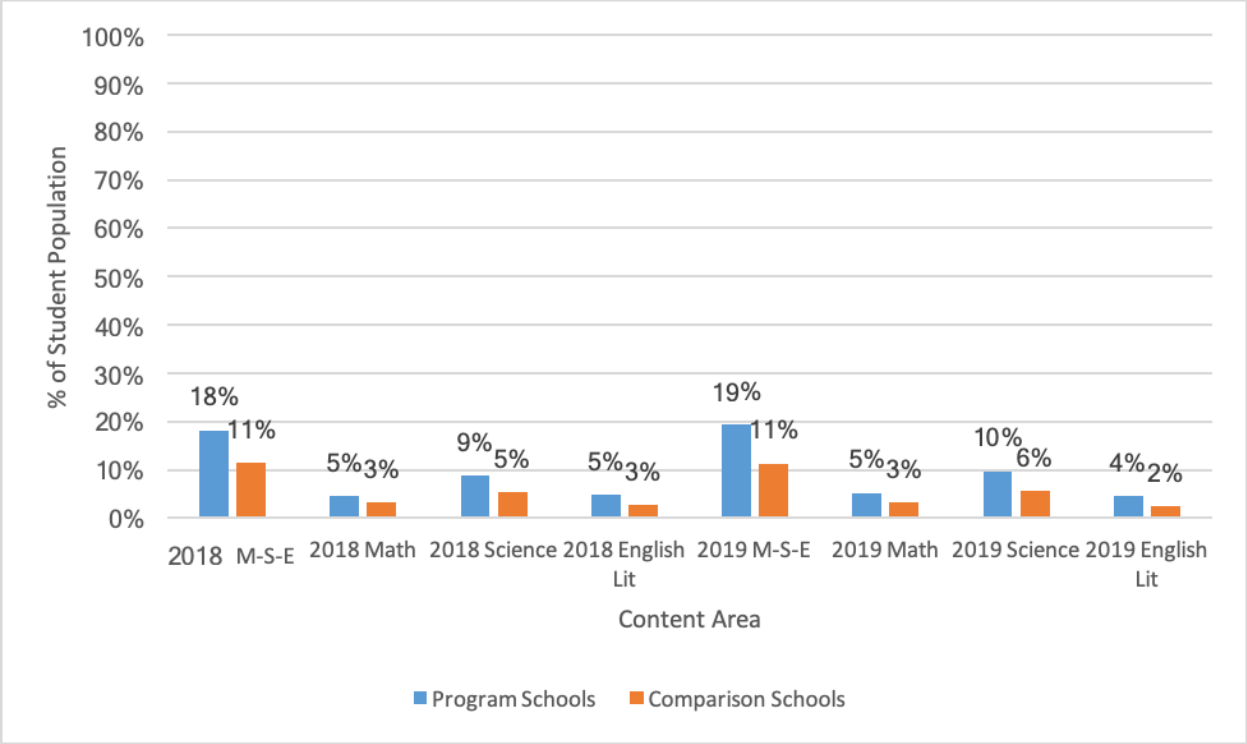



Figure 2. Exams as a percentage of student population.

*Research Question 2: Qualifying scores as a percentage of school population:* For the second outcome, in 2018 the proportion of qualifying scores achieved in the Program Schools was 4.9%, compared to 5.4% in the comparison group. However, this trend was reversed in 2019 as the proportion increased to 5.2% in the Program Schools while it decreased to 5.2% in the Comparison Schools.

*Specific subject areas:* In ELA and math in both years, the percentage of qualifying scores was slightly higher in the Program School populations than in the Comparison Schools. In science, however, in both years the Comparison Schools had a higher percentage of qualifying scores than the Program Schools (see Figure 3).

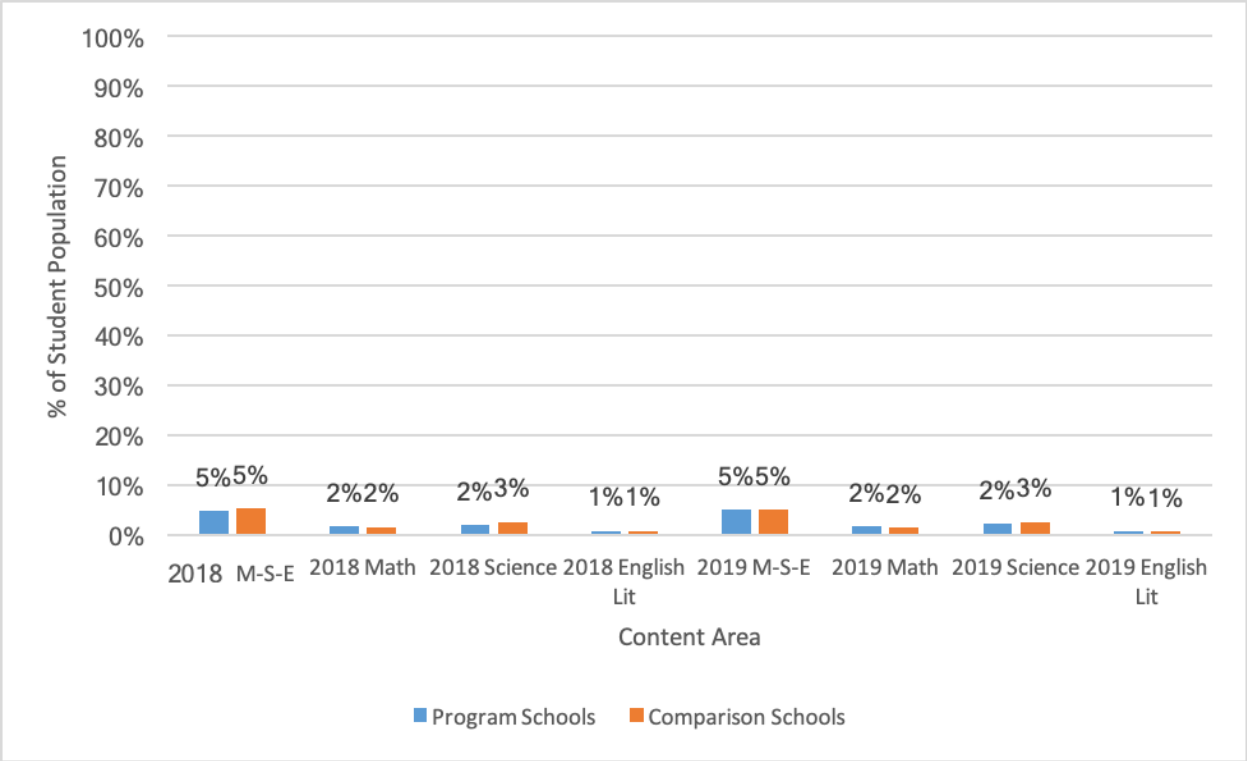



Figure 3. Qualifying scores as a percentage of student population.

*Research Question 3: Qualifying scores as a percentage of tests taken:* When we restrict the sample to students who took AP exams, the difference between the Program and Comparison Schools increases. In 2018, the proportion of qualifying scores among the exam takers was 27% in the Program Schools and 47.2% in the Comparison Schools. In 2019, we see a similar pattern with 26.9% of the program group exams yielding a qualifying score, compared to 45.9% of the comparison group.

*Specific subject areas:* The proportion of exams with qualifying scores was lower in the Program Schools than in the Comparison Schools across all the subjects in both years. The difference was least pronounced in the math 2019 scores (Program School qualifying score rate of 36.6% compared to Comparison School rate of 48.9%) and most pronounced in the 2018 science scores (23.6% for Program Schools and 49.8% for Comparison Schools). See Figure 4.

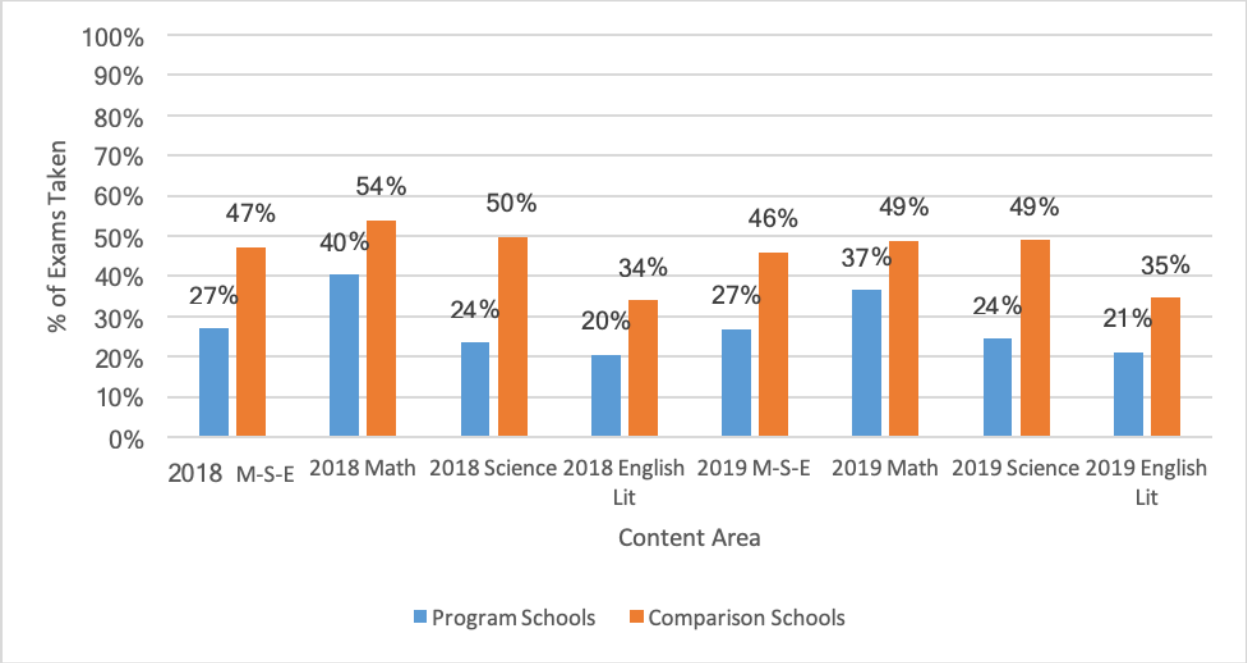



Figure 4. Qualifying scores as a percentage of exams.

Table 2

*Descriptive Statistics of Outcomes in 2018 and 2019 AP Exam Data: Total Exams Taken, Qualifying Scores, Qualifying Scores (Exams Taken)*

| | | 2018 | | | | 2019 | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Treatment | | Comparison | | Treatment | | Comparison | |
| **Subject** | **Outcome** | ***n*** | **%** | ***n*** | **%** | ***n*** | **%** | ***n*** | **%** |
| ELA | Total exams | **2,337** | **4.8** | **1,802** | **2.7** | **2,168** | **4.5** | **1,560** | **2.4** |
| | Qualifying score | 478 | 1.0 | 614 | 0.9 | 457 | 1.0 | 541 | 0.8 |
| | Qualifying score (exams taken) | | 20.5 | | 34.1 | | 21.1 | | 34.7 |
| Math | Total exams | **2,230** | **4.6** | **2,148** | **3.3** | **2,506** | **5.2** | **2,125** | **3.2** |
| | Qualifying score | 901 | 1.9 | 1,158 | 1.8 | 918 | 1.9 | 1,039 | 1.6 |
| | Qualifying score (exams taken) | | 40.4 | | 53.9 | | 36.6 | | 48.9 |
| Science | Total exams | **4,211** | **8.7** | **3,555** | **5.4** | **4,704** | **9.7** | **3,737** | **5.7** |
| | Qualifying score | 993 | 2.1 | 1,769 | 2.7 | 1,148 | 2.4 | 1,828 | 2.8 |
| | Qualifying score (exams taken) | | 23.6 | | 49.8 | | 24.4 | | 48.9 |
| Total | Total exams | **8,778** | **18.1** | **7,505** | **11.3** | **9,378** | **19.4** | **7,422** | **11.2** |
| | Qualifying score | 2,372 | 4.9 | 3,541 | 5.4 | 2,523 | 5.2 | 3,408 | 5.2 |
| | Qualifying score (exams taken) | | 27.0 | | 47.2 | | 26.9 | | 45.9 |

## B. Likelihood of Students Taking an AP Exam or Achieving a Qualifying Score

In order to accommodate different types of clustering of the data, two sets of HGLM analysis were conducted as described in Appendix B. The first set of analyses employed a standard two-level HGLM where students are nested within schools and a treatment indicator variable used as a key predictor in a level 2 model. The second set of analyses used a matched pair analysis where matched pairs of Program and Comparison Schools were considered as a nesting unit and a treatment indicator was included as a key predictor in a level 1 model.

Results of the HGLM analysis of AP exams taken in 2018 are shown in Table 3. The difference between the total number of exams taken across all subjects ("Total exams") at the Program and Comparison Schools was statistically significant. The estimated difference in logit scale between the Program and Comparison Schools was 2.53, with about a 15% difference in estimated probability. The results from the matched pair analysis are in keeping with this result.

The probability of taking an exam in the Program Schools was, on average, 7% higher than the paired Comparison Schools, and the difference was statistically significant.

We next looked at the subject area categories. There was a significantly higher likelihood of taking AP exams at the Program Schools than the Comparison Schools in all subject areas. Specifically, for both ELA and math AP exams, the estimated probability difference between the two groups was about 2%, and the difference in science was 6%.

For the second outcome, qualifying scores within the school population ("Qualifying score"), the standard HGLM analysis indicated no significant difference in the probability of a qualifying score (estimate = 1.24, *p* value = .09, estimated probability difference = 0.01). However, the difference in the matched pair analysis was significant although the estimated difference in probability was very small (1% point). The discrepancy in the result can be explained by the fact that the matched pair analysis carries higher statistical power than the standard HGLM analysis. When taking a closer look at each subject area, science was the only subject in which there was a significant difference between the Program and Comparison Schools for both sets of analyses. However, the direction of the effect was the opposite. The school analysis favored the Program Schools whereas the matched pair analysis showed the comparison condition had a higher probability of achieving a qualifying score compared to the program group.

The third outcome, qualifying scores compared to the number of exams taken ("Qualifying score (exams taken)"), examined whether the probability differed between the two groups. Overall, the fitted probability of achieving a qualifying score among the exams taken was 8% in the Program Schools, compared to 22% in the Comparison Schools. The matched pair analysis also supports this finding (11% for the Program Schools and 21% for Comparison Schools).

For each subject domain, there were no significant differences between Program and Comparison Schools in the school analysis. However, in the matched pair analysis, the fitted probability of achieving a qualifying score on an AP exam at the Program Schools was much lower than Comparison Schools across all subjects. The largest difference was in science (12%).

Table 3

*2-Level HGLM Results: 2018 AP Exam Data*

| Outcome | Var. | Standard 2-level HGLM Analysis | | | | Matched Pair Analysis | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | **Est.** | **S.E.** | ***p* value** | **Fitted prob.** | **Est.** | **S.E.** | ***p* value** | **Fitted prob.** |
| ELA | | | | | | | | | |
| Total exams | Cmp.Int | -5.83 | 0.42 | 0.00 | 0.00 | -4.31 | 0.25 | 0.00 | 0.01 |
| | Diff. (Cmp. Vs. Trt) | 2.14 | 0.55 | 0.00 | 0.02 | 0.64 | 0.03 | 0.00 | 0.02 |
| Qualifying score | Cmp.Int | -8.73 | 0.87 | 0.00 | 0.00 | -7.29 | 0.64 | 0.00 | 0.00 |
| | Diff. (Cmp. Vs. Trt) | 1.22 | 0.94 | 0.19 | 0.00 | 0.05 | 0.06 | 0.43 | 0.00 |
| Qualifying score (exams taken) | Cmp.Int | -2.12 | 0.68 | 0.00 | 0.11 | -2.00 | 0.01 | 0.00 | 0.12 |
| | Diff. (Cmp. Vs. Trt) | -1.03 | 0.86 | 0.23 | 0.04 | -1.09 | 0.01 | 0.00 | 0.04 |
| Math | | | | | | | | | |
| Total exams | Cmp.Int | -5.57 | 0.40 | 0.00 | 0.00 | -4.07 | 0.22 | 0.00 | 0.02 |
| | Diff. (Cmp. Vs. Trt) | 1.76 | 0.53 | 0.00 | 0.02 | 0.38 | 0.03 | 0.00 | 0.02 |
| Qualifying score | Cmp.Int | -7.69 | 0.69 | 0.00 | 0.00 | -6.08 | 0.51 | 0.00 | 0.00 |
| | Diff. (Cmp. Vs. Trt) | 1.43 | 0.83 | 0.09 | 0.00 | 0.04 | 0.05 | 0.35 | 0.00 |
| Qualifying score (exams taken) | Cmp.Int | -1.06 | 0.53 | 0.05 | 0.26 | -1.03 | 0.41 | 0.01 | 0.26 |
| | Diff. (Cmp. Vs. Trt) | -0.42 | 0.69 | 0.54 | 0.19 | -0.42 | 0.09 | 0.00 | 0.19 |
| Science | | | | | | | | | |
| Total exams | Cmp.Int | -5.0 | 0.36 | 0.00 | 0.01 | -3.21 | 0.14 | 0.00 | 0.04 |
| | Diff. (Cmp. Vs. Trt) | 2.39 | 0.48 | 0.00 | 0.07 | 0.55 | 0.02 | 0.00 | 0.07 |
| Qualifying score | Cmp.Int | -7.94 | 0.71 | 0.00 | 0.00 | -5.35 | 0.40 | 0.00 | 0.01 |
| | Diff. (Cmp. Vs. Trt) | 2.03 | 0.85 | 0.02 | 0.00 | -0.31 | 0.04 | 0.00 | 0.00 |
| Qualifying score (exams taken) | Cmp.Int | -2.10 | 0.58 | 0.00 | 0.11 | -1.33 | 0.35 | 0.00 | 0.21 |
| | Diff. (Cmp. Vs. Trt) | -0.53 | 0.72 | 0.46 | 0.07 | -1.04 | 0.07 | 0.00 | 0.09 |
| All subjects | | | | | | | | | |
| Total exams | Cmp.Int | -4.09 | 0.35 | 0.00 | 0.02 | -2.39 | 0.15 | 0.00 | 0.08 |
| | Diff. (Cmp. Vs. Trt) | 2.53 | 0.47 | 0.00 | 0.17 | 0.63 | 0.02 | 0.00 | 0.15 |
| Qualifying score | Cmp.Int | -6.09 | 0.56 | 0.00 | 0.00 | -4.49 | 0.40 | 0.00 | 0.01 |
| | Diff. (Cmp. Vs. Trt) | 1.24 | 0.74 | 0.09 | 0.01 | -0.13 | 0.03 | 0.00 | 0.01 |
| Qualifying score (exams taken) | Cmp.Int | -1.24 | 0.44 | 0.01 | 0.22 | -1.33 | 0.33 | 0.00 | 0.21 |
| | Diff. (Cmp. Vs.Trt) | -1.22 | 0.58 | 0.03 | 0.08 | -0.79 | 0.04 | 0.00 | 0.11 |

*Note.* The estimated probability for Diff. (Cmp. Vs. Trt) in the table is the fitted probability for Program Schools.

The HGLM results using the 2019 school-year AP exam data are shown in Table 4. For most cases, the results were similar to those from the previous year. In total, the fitted probability of taking an AP exam was significantly higher in Program Schools (18%) than in Comparison Schools (3%). The significant and positive treatment effect was retained in the matched pair analysis, with a difference of 8% between the two groups of schools. Across all subject domains the differences between the treatment and comparison conditions were significant and higher in the Program Schools. The difference between the two groups was the largest in science at 7%, and about 2% in English and math.

For the second outcome, qualifying scores within the school population, exams taken at the Program Schools had a significantly higher overall probability (2%) of receiving a qualifying score than the Comparison Schools (0%). Unlike the corresponding results from the previous year, there was no significant difference in the matched pair analysis.

For the different subject-area domains, the fitted probability of the program group to achieve a qualifying score on AP math exams was significantly higher than the comparison condition in both the school and the matched pair analyses. However, the fitted difference between the two groups in math was very small on the probability scale. In contrast, the program group exams had a significantly lower chance of obtaining a qualifying score in science than the comparison in matched pair analysis, with a difference of 1%. For ELA the difference between the Program and Comparison Schools was not statistically significant.

Results of the standard HGLM analysis for the third outcome, qualifying scores among exams taken, showed that the difference between the Program Schools (7%) and the Comparison Schools (9%) was not statistically significant.

In total, the fitted probability of the treatment condition to have a higher percentage of qualifying scores among exams taken was significantly lower than the comparison group by 10% points in the matched pair analysis. This pattern was consistent across all subject area domains. The largest difference between the program and comparison groups was an 11% difference in science exams, followed by a 10% difference in math, and a 5% difference in ELA.

Table 4

*2-Level HGLM Results: 2019 AP Exam Data*

| Outcome | Var. | Standard 2-level HGLM Analysis | | | | Matched Pair Analysis | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Est. | S.E. | *p* value | Fitted prob. | Est. | S.E. | *p* value | Fitted prob. |
| ELA | | | | | | | | | |
| Total exams | Cmp.Int | -5.74 | 0.39 | 0.00 | 0.00 | -4.44 | 0.24 | 0.00 | 0.01 |
| | Diff. (Cmp. Vs. Trt) | 1.98 | 0.51 | 0.00 | 0.02 | 0.70 | 0.03 | 0.00 | 0.02 |
| Qualifying score | Cmp.Int | -10.41 | 1.18 | 0.00 | 0.00 | -7.35 | 0.61 | 0.00 | 0.00 |
| | Diff. (Cmp. Vs. Trt) | 2.01 | 1.16 | 0.08 | 0.00 | 0.12 | 0.06 | 0.06 | 0.00 |
| Qualifying score (exams taken) | Cmp.Int | -3.39 | 0.88 | 0.00 | 0.03 | -2.18 | 0.50 | 0.00 | 0.10 |
| | Diff. (Cmp. Vs. Trt) | -0.09 | 1.03 | 0.93 | 0.03 | -0.68 | 0.10 | 0.00 | 0.05 |
| Math | | | | | | | | | |
| Total exams | Cmp.Int | -5.44 | 0.40 | 0.00 | 0.00 | -4.07 | 0.23 | 0.00 | 0.02 |
| | Diff. (Cmp. Vs. Trt) | 1.63 | 0.54 | 0.00 | 0.02 | 0.52 | 0.03 | 0.00 | 0.03 |
| Qualifying score | Cmp.Int | -7.75 | 0.66 | 0.00 | 0.00 | -5.86 | 0.44 | 0.00 | 0.00 |
| | Diff. (Cmp. Vs. Trt) | 1.80 | 0.79 | 0.02 | 0.00 | 0.17 | 0.05 | 0.00 | 0.00 |
| Qualifying score (exams taken) | Cmp.Int | -1.50 | 0.51 | 0.00 | 0.18 | -0.90 | 0.35 | 0.01 | 0.29 |
| | Diff. (Cmp. Vs. Trt) | 0.14 | 0.65 | 0.83 | 0.20 | -0.54 | 0.08 | 0.00 | 0.19 |
| Science | | | | | | | | | |
| Total exams | Cmp.Int | -4.57 | 0.32 | 0.00 | 0.01 | -3.14 | 0.15 | 0.00 | 0.04 |
| | Diff. (Cmp. Vs. Trt) | 2.14 | 0.43 | 0.00 | 0.08 | 0.63 | 0.02 | 0.00 | 0.08 |
| Qualifying score | Cmp.Int | -7.38 | 0.67 | 0.00 | 0.00 | -5.38 | 0.43 | 0.00 | 0.01 |
| | Diff. (Cmp. Vs. Trt) | 1.50 | 0.82 | 0.07 | 0.00 | -0.18 | 0.04 | 0.00 | 0.00 |
| Qualifying score (exams taken) | Cmp.Int | -1.87 | 0.55 | 0.00 | 0.13 | -1.42 | 0.37 | 0.00 | 0.19 |
| | Diff. (Cmp. Vs. Trt) | -0.92 | 0.69 | 0.18 | 0.06 | -1.06 | 0.07 | 0.00 | 0.08 |
| All subjects | | | | | | | | | |
| Total exams | Cmp.Int | -3.64 | 0.30 | 0.00 | 0.03 | -2.38 | 0.14 | 0.00 | 0.08 |
| | Diff. (Cmp. Vs. Trt) | 2.14 | 0.41 | 0.00 | 0.18 | 0.73 | 0.02 | 0.00 | 0.16 |
| Qualifying score | Cmp.Int | -6.75 | 0.65 | 0.00 | 0.00 | -4.42 | 0.38 | 0.00 | 0.01 |
| | Diff. (Cmp. Vs. Trt) | 1.92 | 0.83 | 0.02 | 0.01 | -0.01 | 0.03 | 0.81 | 0.01 |
| Qualifying score (exams taken) | Cmp.Int | -2.27 | 0.52 | 0.00 | 0.09 | -1.35 | 0.31 | 0.00 | 0.21 |
| | Diff. (Cmp. Vs. Trt) | -0.32 | 0.67 | 0.64 | 0.07 | -0.76 | 0.04 | 0.00 | 0.11 |

*Note.* The estimated probability for Diff. (Cmp. Vs. Trt) in the table is the fitted probability for Program Schools.

## C. Likelihood of Students Taking an AP Exam or Achieving a Qualifying Score Controlling for School Size

In addition to the unconditional HGLM analyses, conditional HGLMs were fitted to control for the possible confounding of treatment condition with school size and to obtain more precise estimates of program impact. For the conditional (or standard) HGLM analysis, the covariate was the number of students in each school. For the matched pair analysis, the average number of students in each pair was used as a covariate. In order to avoid nonconvergence, a scaled version of the school size was included in the analysis models, which was computed by $Size_j = \frac{School\ size_j - \overline{School\ size}}{SD(School\ size)}$.

In conditional models, the intercept means the fitted probability in the Comparison Schools where the number of students equals the average across schools and the treatment slope is interpreted as the expected difference in probability between the Comparison and Program Schools, holding the school size (or the average school size within a matched pair) constant.

For the 2018 AP exam data, most of the results from the standard (conditional) HGLM analysis with school size as a covariate agreed with the results from the corresponding unconditional models except for the third outcome, qualifying scores among exams taken. Since the Program Schools were matched with the Comparison Schools based on school characteristics including school size, the conditional model results from the matched pair analyses were almost the same with those from the null models, as expected (see Table 5).

For the first outcome, "Total exams," in the conditional HGLM analysis the probability of taking an AP exam after controlling for school size was 20% in the Program Schools, which was significantly higher than in the Comparison Schools (which had a 2% probability). The results from the matched pair analysis show a similar pattern, albeit with a narrower gap of 7% between the two groups. For the specific subject areas, in the conditional HGLM analysis, the expected difference between Program and Comparison Schools were all significant as in the unconditional model results.

The second outcome, qualifying score based on school population, was also significant in the conditional HGLM analysis with a higher fitted probability of receiving qualifying scores on AP exams in Program Schools (1%) than in Comparison Schools (0%), after controlling for school size. In terms of subject domains (ELA, math, and science), the only difference from the unconditional model results was found in math where the estimated treatment slope was 1.80 in logit scale and significant. For the other subject areas in the standard HGLM analysis and all

subject areas in matched pair analysis, the results remain unchanged when compared to the HGLM results without any covariate, with negligible difference in probability scale.

The analysis results of the last outcome, “Qualifying score (exams taken),” show that the fitted probability of receiving a qualifying score among the exams taken was higher in the Comparison Schools as was the case in the unconditional model results. However, the difference between the two groups was not significant with a gap narrowed down from 14% to 10%. For the specific subject domains, the results were similar to those from the unconditional models, except for science in the conditional HGLM analysis. In science, the difference between the two conditions became significant when the school size was included in the model, with the estimated probability of 9% in Program Schools and 10% in Comparison Schools. The results from the unconditional and conditional models supported the results of the matched pair analyses.

Table 5

*2-Level HGLM Results Controlling for School Size: 2018 AP Exam Data*

| Outcome | Var. | Standard 2-level HGLM Analysis | | | | Matched Pair Analysis | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | **Est.** | **S.E.** | ***p* value** | **Fitted prob.** | **Est.** | **S.E.** | ***p* value** | **Fitted prob.** |
| ELA | | | | | | | | | |
| Total exams | Cmp.Int | -5.78 | 0.42 | 0.00 | 0.00 | -4.24 | 0.27 | 0.00 | 0.01 |
| | Diff. (Cmp. Vs. Trt) | 2.38 | 0.59 | 0.00 | 0.03 | 0.64 | 0.03 | 0.00 | 0.03 |
| | Size | 0.38 | 0.31 | 0.22 | | 0.17 | 0.27 | 0.52 | |
| Qualifying score | Cmp.Int | -8.65 | 0.84 | 0.00 | 0.00 | -7.02 | 0.66 | 0.00 | 0.00 |
| | Diff. (Cmp. Vs. Trt) | 1.66 | 1.00 | 0.10 | 0.00 | 0.05 | 0.06 | 0.43 | 0.00 |
| | Size | 0.74 | 0.54 | 0.17 | | 0.71 | 0.63 | 0.26 | |
| Qualifying score (exams taken) | Cmp.Int | -2.24 | 0.70 | 0.00 | 0.10 | -1.94 | 0.54 | 0.00 | 0.13 |
| | Diff. (Cmp. Vs. Trt) | -0.77 | 0.93 | 0.41 | 0.05 | -1.09 | 0.10 | 0.00 | 0.05 |
| | Size | 0.35 | 0.49 | 0.48 | | 0.40 | 0.53 | 0.45 | |
| Math | | | | | | | | | |
| Total exams | Cmp.Int | -5.51 | 0.40 | 0.00 | 0.00 | -3.94 | 0.24 | 0.00 | 0.02 |
| | Diff. (Cmp. Vs. Trt) | 2.08 | 0.57 | 0.00 | 0.03 | 0.38 | 0.03 | 0.00 | 0.03 |
| | Size | 0.52 | 0.30 | 0.08 | | 0.29 | 0.24 | 0.23 | |
| Qualifying score | Cmp.Int | -7.60 | 0.69 | 0.00 | 0.00 | -5.77 | 0.53 | 0.00 | 0.00 |
| | Diff. (Cmp. Vs. Trt) | 1.80 | 0.88 | 0.04 | 0.00 | 0.04 | 0.05 | 0.36 | 0.00 |
| | Size | 0.67 | 0.48 | 0.16 | | 0.76 | 0.53 | 0.15 | |
| Qualifying score (exams taken) | Cmp.Int | -1.10 | 0.54 | 0.04 | 0.25 | -0.93 | 0.42 | 0.03 | 0.28 |
| | Diff. (Cmp. Vs. Trt) | -0.27 | 0.74 | 0.71 | 0.20 | -0.42 | 0.09 | 0.00 | 0.21 |
| | Size | 0.21 | 0.40 | 0.60 | | 0.41 | 0.40 | 0.31 | |

| Outcome | Var. | Standard 2-level HGLM Analysis | | | | Matched Pair Analysis | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Est. | S.E. | *p* value | Fitted prob. | Est. | S.E. | *p* value | Fitted prob. |
| Science | | | | | | | | | |
| Total exams | Cmp.Int | -4.97 | 0.36 | 0.00 | 0.01 | -3.21 | 0.16 | 0.00 | 0.04 |
| | Diff. (Cmp. Vs. Trt) | 2.58 | 0.51 | 0.00 | 0.08 | 0.55 | 0.02 | 0.00 | 0.07 |
| | Size | 0.31 | 0.27 | 0.25 | | 0.02 | 0.15 | 0.89 | |
| Qualifying score | Cmp.Int | -7.84 | 0.70 | 0.00 | 0.00 | -4.94 | 0.41 | 0.00 | 0.01 |
| | Diff. (Cmp. Vs. Trt) | 2.58 | 0.91 | 0.00 | 0.01 | -0.31 | 0.04 | 0.00 | 0.01 |
| | Size | 0.93 | 0.48 | 0.06 | | 1.03 | 0.42 | 0.01 | |
| Qualifying score (exams taken) | Cmp.Int | -2.24 | 0.00 | 0.00 | 0.10 | -1.07 | 0.34 | 0.00 | 0.25 |
| | Diff. (Cmp. Vs. Trt) | -0.03 | 0.00 | 0.00 | 0.09 | -1.03 | 0.07 | 0.00 | 0.11 |
| | Size | 0.70 | 0.00 | 0.00 | | 0.89 | 0.34 | 0.01 | |
| All subjects | | | | | | | | | |
| Total exams | Cmp.Int | -4.05 | 0.35 | 0.00 | 0.02 | -2.40 | 0.16 | 0.00 | 0.08 |
| | Diff. (Cmp. Vs. Trt) | 2.66 | 0.50 | 0.00 | 0.20 | 0.63 | 0.02 | 0.00 | 0.15 |
| | Size | 0.22 | 0.26 | 0.40 | | -0.02 | 0.16 | 0.92 | |
| Qualifying score | Cmp.Int | -5.97 | 0.56 | 0.00 | 0.00 | -4.12 | 0.41 | 0.00 | 0.02 |
| | Diff. (Cmp. Vs. Trt) | 1.70 | 0.78 | 0.03 | 0.01 | -0.13 | 0.03 | 0.00 | 0.01 |
| | Size | 0.82 | 0.42 | 0.05 | | 0.88 | 0.42 | 0.04 | |
| Qualifying score (exams taken) | Cmp.Int | -1.35 | 0.44 | 0.00 | 0.21 | -1.08 | 0.32 | 0.00 | 0.25 |
| | Diff. (Cmp. Vs. Trt) | -0.77 | 0.61 | 0.21 | 0.11 | -0.79 | 0.04 | 0.00 | 0.13 |
| | Size | 0.63 | 0.33 | 0.06 | | 0.78 | 0.32 | 0.01 | |

*Note.* The estimated probability for Diff. (Cmp. Vs. Trt) in the table is the fitted probability for Program Schools.

Table 6 shows the conditional HGLM analysis results for the 2019 school-year AP exam data using school size as a covariate. The 2019 results support the corresponding results from the unconditional models. The only discrepancy was the second outcome, qualifying scores based on school population, for ELA and science in the standard HGLM analysis.

The analysis of the first outcome, “Total exams,” revealed that the probability of taking an AP exam is higher in Program Schools compared to the Comparison Schools. The results were the same as those from the unconditional models. All the subject domains showed a similar pattern to the HGLM analysis without any covariates, with a higher probability of taking an AP exam in the Program Schools.

For the second outcome, qualifying scores based on school population, exams taken at Program Schools had a higher fitted probability of resulting in a qualifying score than in Comparison Schools. The difference was significant in the conditional HGLM analysis, while it was not significant in the matched pair analysis. The results correspond to those from the unconditional models, with negligible differences in estimated probabilities. For the specific subject area domains, however, the expected difference which favored the Program Schools became significant for English and science, though the fitted probability for each group remained almost the same. The conditional HGLM analysis result for math and the matched pair analysis results for all subjects remain unchanged from the unconditional models.

For the third outcome, qualifying scores among exams taken, the expected difference between the Program and Comparison Schools in total was only significant in the matched pair analysis. Overall, the fitted probability of achieving a qualifying score among the exams taken was higher in the comparison group (24%) than in the Program Schools (13%). These results are consistent with those from the unconditional models. For the specific subject domains, the expected differences were not significant in all domains in the school analysis whereas the fitted probability in the Comparison Schools was significantly higher in all subjects in the matched pair analysis. The largest gap between the two groups was found in science, which was 13%, followed by 10% in math, and the smallest difference was 5% in English.

Table 6

*2-Level HGLM Results Controlling for School Size: 2019 AP Exam Data*

| Outcome | Var. | Standard 2-level HGLM Analysis | | | | Matched Pair Analysis | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Est. | S.E. | *p* value | Fitted prob. | Est. | S.E. | *p* value | Fitted prob. |
| ELA | | | | | | | | | |
| Total exams | Cmp.Int | -5.69 | 0.39 | 0.00 | 0.00 | -4.36 | 0.26 | 0.00 | 0.01 |
| | Diff. (Cmp. Vs. Trt) | 2.17 | 0.54 | 0.00 | 0.03 | 0.70 | 0.03 | 0.00 | 0.03 |
| | Size | 0.31 | 0.28 | 0.27 | | 0.20 | 0.25 | 0.42 | |
| Qualifying score | Cmp.Int | -10.10 | 1.12 | 0.00 | 0.00 | -7.08 | 0.63 | 0.00 | 0.00 |
| | Diff. (Cmp. Vs. Trt) | 2.50 | 1.21 | 0.04 | 0.00 | 0.12 | 0.06 | 0.06 | 0.00 |
| | Size | 0.90 | 0.63 | 0.15 | | 0.73 | 0.59 | 0.22 | |
| Qualifying score (exams taken) | Cmp.Int | -3.57 | 0.89 | 0.00 | 0.03 | -2.10 | 0.50 | 0.00 | 0.11 |
| | Diff. (Cmp. Vs. Trt) | 0.39 | 1.10 | 0.72 | 0.04 | -0.68 | 0.10 | 0.00 | 0.06 |
| | Size | 0.73 | 0.58 | 0.21 | | 0.45 | 0.49 | 0.36 | |
| Math | | | | | | | | | |
| Total exams | Cmp.Int | -5.36 | 0.40 | 0.00 | 0.00 | -3.96 | 0.25 | 0.00 | 0.02 |
| | Diff. (Cmp. Vs. Trt) | 2.00 | 0.57 | 0.00 | 0.03 | 0.52 | 0.03 | 0.00 | 0.03 |
| | Size | 0.59 | 0.29 | 0.04 | | 0.28 | 0.24 | 0.25 | |
| Qualifying score | Cmp.Int | -7.67 | 0.65 | 0.00 | 0.00 | -5.64 | 0.46 | 0.00 | 0.00 |
| | Diff. (Cmp. Vs. Trt) | 2.23 | 0.84 | 0.01 | 0.00 | 0.17 | 0.05 | 0.00 | 0.00 |
| | Size | 0.69 | 0.44 | 0.11 | | 0.55 | 0.45 | 0.23 | |
| Qualifying score (exams taken) | Cmp.Int | -1.54 | 0.51 | 0.00 | 0.18 | -0.85 | 0.36 | 0.02 | 0.30 |
| | Diff. (Cmp. Vs. Trt) | 0.26 | 0.70 | 0.70 | 0.22 | -0.54 | 0.08 | 0.00 | 0.20 |
| | Size | 0.17 | 0.36 | 0.64 | | 0.21 | 0.34 | 0.54 | |

| Outcome | Var. | Standard 2-level HGLM Analysis | | | | Matched Pair Analysis | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Est. | S.E. | *p* value | Fitted prob. | Est. | S.E. | *p* value | Fitted prob. |
| Science | | | | | | | | | |
| Total exams | Cmp.Int | -4.54 | 0.32 | 0.00 | 0.01 | -3.16 | 0.16 | 0.00 | 0.04 |
| | Diff. (Cmp. Vs. Trt) | 2.25 | 0.45 | 0.00 | 0.09 | 0.63 | 0.02 | 0.00 | 0.07 |
| | Size | 0.17 | 0.23 | 0.45 | | -0.03 | 0.16 | 0.86 | |
| Qualifying score | Cmp.Int | -7.28 | 0.66 | 0.00 | 0.00 | -5.11 | 0.45 | 0.00 | 0.01 |
| | Diff. (Cmp. Vs. Trt) | 1.99 | 0.87 | 0.02 | 0.01 | -0.18 | 0.04 | 0.00 | 0.01 |
| | Size | 0.83 | 0.46 | 0.07 | | 0.67 | 0.44 | 0.13 | |
| Qualifying score (exams taken) | Cmp.Int | -1.99 | 0.55 | 0.00 | 0.12 | -1.26 | 0.37 | 0.00 | 0.22 |
| | Diff. (Cmp. Vs. Trt) | -0.49 | 0.73 | 0.51 | 0.08 | -1.06 | 0.07 | 0.00 | 0.09 |
| | Size | 0.64 | 0.38 | 0.09 | | 0.62 | 0.36 | 0.08 | |
| All subjects | | | | | | | | | |
| Total exams | Cmp.Int | -3.62 | 0.30 | 0.00 | 0.03 | -2.40 | 0.16 | 0.00 | 0.08 |
| | Diff. (Cmp. Vs. Trt) | 2.20 | 0.44 | 0.00 | 0.19 | 0.73 | 0.02 | 0.00 | 0.16 |
| | Size | 0.10 | 0.22 | 0.66 | | -0.04 | 0.15 | 0.82 | |
| Qualifying score | Cmp.Int | -6.64 | 0.64 | 0.00 | 0.00 | -4.18 | 0.40 | 0.00 | 0.02 |
| | Diff. (Cmp. Vs. Trt) | 2.40 | 0.87 | 0.01 | 0.01 | -0.01 | 0.03 | 0.81 | 0.02 |
| | Size | 0.80 | 0.45 | 0.07 | | 0.58 | 0.40 | 0.14 | |
| Qualifying score (exams taken) | Cmp.Int | -2.36 | 0.52 | 0.00 | 0.09 | -1.18 | 0.31 | 0.00 | 0.24 |
| | Diff. (Cmp. Vs. Trt) | 0.16 | 0.70 | 0.82 | 0.10 | -0.76 | 0.04 | 0.00 | 0.13 |
| | Size | 0.71 | 0.36 | 0.05 | | 0.55 | 0.30 | 0.07 | |

*Note.* The estimated probability for Diff. (Cmp. Vs. Trt) in the table is the fitted probability for Program Schools.

## D. Descriptive Analysis STEM Major Enrollment Data

The basis for this descriptive analysis was college STEM major enrollment data obtained from the National Student Clearing House (NSC), and high school enrollment data obtained from the National Center for Education Statistics (NCES). Data from the NSC were available for the Treatment and Delayed Treatment School students. The first group of 27 schools first participated in the CRP in the 2016–2017 school year, and the NSC was able to provide college STEM major enrollment data for students in these 27 schools who graduated from high school in 2017, 2018, or 2019. The second group of 21 schools (the delayed treatment group) first participated in the CRP in the 2017–2018 school year. The NSC was able to provide college STEM major enrollment data for students in this group who graduated high school in either 2018

or 2019. The current analyses looked to answer the following question: To what extent does the percentage of students enrolled in STEM majors in college vary across years following treatment?

We first linked NSC data to NCES data to determine the number of students enrolled in STEM majors for each school for each year of graduation. The STEM major enrollment counts were next linked to the school population data in the appropriate year for the 48 schools in the analysis. The percentage of students in each school who ultimately enrolled in STEM majors in each school and year of graduation was calculated by dividing the STEM major enrollment counts by the 12th grade student population in the appropriate year, and then multiplying by 100.

The average percentage of CRP students who later enrolled in a STEM major in college is presented for the two groups of Treatment Schools in Table 7 (grouped by year of graduation).

Table 7

*Average Percentage of Students Who Enroll in a STEM Major*

| Treatment type | Graduation year | STEM major enrollment | | | | |
|---|---|---|---|---|---|---|
| | | ***n*** | **Min (%)** | **Max (%)** | ***M* (%)** | ***SD*** |
| Treatment | 2016–2017 | 25 | 0.00 | 55.06 | 14.80 | 11.76 |
| | 2017–2018 | 27 | 2.59 | 46.86 | 17.41 | 11.09 |
| | 2018–2019 | 27 | 0.65 | 29.45 | 14.79 | 7.96 |
| Delayed Treatment | 2017–2018 | 21 | 0.00 | 37.59 | 17.06 | 11.23 |
| | 2018–2019 | 21 | 0.30 | 34.93 | 17.05 | 11.48 |
| All Treatment Schools | 2016–2017 | 25 | 0.00 | 55.06 | 14.80 | 11.76 |
| | 2017–2018 | 48 | 0.00 | 46.86 | 17.26 | 11.04 |
| | 2018–2019 | 48 | 0.30 | 34.93 | 15.78 | 9.61 |

*Note.* Delayed Treatment Schools began treatment in 2017–2018 and not included in the descriptive results for 2016–2017.

For the group of schools that began treatment in 2016–2017, the average percentage of STEM majors displayed a modest increase from 14.8 % for the students who graduated in 2017, to 17.41% for those who graduated in 2018. For the group who graduated in 2019, there was a decrease in the average percentage back to the 2017 level (14.79%).

For the Delayed Treatment group, the average percentage of 12th grade students who later enrolled in a college STEM major was almost identical in 2017–2018 ($M$ = 17.05, $SD$ = 11.23) and 2018–2019 ($M$ = 17.06, $SD$ = 11.48). As can be seen in Table 7, variance in the percentage

of students who enrolled in a STEM college major was high with a standard deviation of ~11% in both years.

### E. HGLM (NMSI Schools only using implementation variables)

In this section, we present the basic population samples and methods used to examine the relationship between the CRP implementation indicators (drawn from the fidelity matrix analyses) and student outcomes using the 2017–2018 and 2018–2019 AP exam data. For this analysis, we included the NMSI schools where the school-level information was available (i.e., the 48 program schools). As described elsewhere, the two treatment implementation groups reflect differing lengths of implementation (either two or three years).

The two treatment implementation groups were compared on three primary outcome measures. The relationship between the length of implementation and each outcome measure informs research questions as presented in section III of this report which concerned the likelihood of students either taking AP exams or receiving a qualifying score on an AP exam. The outcome measures were examined across ELA, math, and science subjects individually, and then for the three subjects grouped together ("All subjects") for both the 2017–2018 and 2018–2019 school years. The outcome measures were specifically defined to inform the research questions described above.

1. The likelihood of a student taking an AP exam.
2. The likelihood of a student receiving a qualifying score on an AP exam.
3. The likelihood of a student receiving a qualifying score on an AP exam (only for those students who took AP exams).

In order to accommodate the clustering of the student data within schools, we employed a standard two-level Hierarchical General Linear Model (HGLM) where students were nested within schools and a treatment length indicator variable was used as a key predictor in a level 2 model. Since our outcome scores have values of either zero or one, we cannot assume a normal distribution and thus used a logit link function. We included six covariates potentially related to CRP implementation gleaned from school-level information and teacher survey responses. The teacher survey responses were first aggregated at the school level as means or percentages. Covariates were then standardized at the school level in order to make the interpretation meaningful and to avoid nonconvergence. Table 8 provides details of these covariates.

Table 8

*Descriptions of the Variables Used in the HGLM Models*

| Variable | Description | Measure |
|---|---|---|
| Implementation Group | Delayed Treatment / Treatment | Delayed treatment = 0; Treatment = 1 |
| Teaching experience | Number of years as a teacher | Mean |
| AP teaching experience | Number of years as an AP teacher | Mean |
| PSD motivation | Level of agreement that school's PSD motivates teachers and students to see the benefits of AP coursework | Mean |
| Open enrollment impact | Open enrollment positive impact on AP | Mean (0/1) |
| CRP effectiveness | Has CRP made improvements across seven goals | Sum/ Mean |
| Teacher engagement | Number of survey participants | Percentage of teachers |

First, we present the HGLM results with covariates based on 2018 AP data (Table 9). In this table, for each model presented, the fitted probability of the intercept represents the expected probability of the outcome for the Delayed Treatment Schools in which the value of each variable in the model is at the mean of each variable. Thus, it represents the expected probability for the "average" delayed treatment group schools. Similarly, we present the expected probability for the "average" Treatment Schools in the fitted probability column in the result table. Note, however, that estimate of the treatment implementation indicator (Yr 1 vs. Yr 2) is the expected difference in logit between the two different groups of schools. The fitted probabilities for the other covariates are presented only when statistically significant. The fitted probability for a covariate is interpreted as the expected probability of the outcome in the "average" Delayed Treatment School when the covariate changes in one unit. We order our discussion by describing all results across subject categories for each outcome in the sequence.

The first outcome ("Total exams") is the likelihood of taking an AP exam in 2018 data. For "Total exams" across all subjects, the estimated (fitted) probability of taking an AP exam was 13% for the schools where the CRP had been implemented for one year as compared to 23% for the schools where the CRP had been implemented for two years holding all the covariates in the model constant. As the difference indicates, the length of implementation of CRP was significantly and positively associated with "Total exams," holding other predictors constant (Est. = 0.69, S.E. = 0.32, fitted probability = 0.23).

We found a similar result for the likelihood of taking science AP exams. In science, the estimated probability for the Treatment Schools was significantly higher than for the Delayed

Treatment Schools by a fitted probability difference of 5%, controlling for other covariates in the model (Est. = 0.87, S.E. = 0.33, fitted probability = 0.10). The length of treatment implementation indicator, however, was not a statistically significant predictor of the likelihood of taking ELA or math AP exams. In addition, the other six covariates (other than year of treatment implementation) were not statistically significant predictors of the likelihood of students taking an AP exam.

The second outcome, qualifying scores within the school population ("Qualifying score"), had no significant relationship with length of treatment implementation. This was the case when summed over all subjects and for each individual subject category. There were two covariates, however, that were significantly related to the likelihood of achieving qualifying scores. The covariate "CRP effectiveness" was a significant predictor in ELA AP exams (Est. =1.46, S.E. = 0.67, fitted probability = 0.01). Specifically, a unit increase in "CRP effectiveness" was associated with a 1% increase in the fitted probability, holding other predictors constant. Similarly, for science AP exams, "AP teaching experience" had a positive relationship with the outcome, where a unit change in the covariate was associated with about 0% increase in the fitted probability, controlling for other covariates (Est. =0.86, S.E. = 0.44, fitted probability = 0.01).

For the third outcome, qualifying scores compared to the number of exams taken ("Qualifying score (exams taken)"), no covariate had a significant relationship with the likelihood of achieving a qualifying score when the outcome was evaluated across all subjects. When looking at each subject category, however, there were some significant findings. For ELA exams, a unit increase in "Teaching experience" was associated with a 19% increase in the fitted probability (Est. =1.21, S.E. = 0.46, fitted probability = 0.31). In math, the fitted probability in the Treatment Schools was lower than in the Delayed Treatment Schools by 43%, with other covariates held constant (Est. =-2.31, S.E. = 0.86, fitted probability = 0.09). A unit increase in "Teacher engagement" was associated with 28% increase in the fitted probability for math AP exams (Est. =1.30, S.E. = 0.42, fitted probability = 0. 80). In the science domain, the probability for the Treatment Schools was lower by 20% than for the Delayed Treatment Schools (Est. =-1.74, S.E. = 0.87, fitted probability = 0.05).

Table 9

*2-Level HGLM Results: 2018 AP Exam Data*

| Outcome | Variable | Est. | S.E | p value | Fitted prob |
|---|---|---|---|---|---|
| Total exams | | | | | |
| ELA | Intercept | -3.02 | 0.46 | 0.00 | 0.05 |
| | Implementation (Yr1 vs. Yr2) | -0.78 | 0.62 | 0.20 | |
| | Teaching experience | -0.37 | 0.31 | 0.24 | |
| | AP teaching experience | 0.06 | 0.33 | 0.85 | |
| | PSD motivation | -0.13 | 0.32 | 0.69 | |
| | Open enrollment impact | -0.09 | 0.42 | 0.82 | |
| | CRP effectiveness | 0.42 | 0.33 | 0.21 | |
| | Teacher engagement | 0.12 | 0.32 | 0.70 | |
| Math | Intercept | -3.82 | 0.38 | 0.00 | 0.02 |
| | Implementation (Yr1 vs. Yr2) | 0.84 | 0.49 | 0.08 | |
| | Teaching experience | 0.12 | 0.24 | 0.60 | |
| | AP teaching experience | 0.04 | 0.24 | 0.86 | |
| | PSD motivation | 0.20 | 0.24 | 0.41 | |
| | Open enrollment impact | 0.10 | 0.31 | 0.75 | |
| | CRP effectiveness | -0.14 | 0.26 | 0.58 | |
| | Teacher engagement | -0.27 | 0.24 | 0.26 | |
| Science | Intercept | -3.02 | 0.25 | 0.00 | 0.05 |
| | Implementation (Yr1 vs. Yr2) | 0.87 | 0.33 | 0.01 | 0.10 |
| | Teaching experience | 0.03 | 0.17 | 0.88 | |
| | AP teaching experience | 0.29 | 0.17 | 0.08 | |
| | PSD motivation | -0.12 | 0.17 | 0.47 | |
| | Open enrollment impact | 0.00 | 0.21 | 0.99 | |
| | CRP effectiveness | -0.26 | 0.18 | 0.16 | |
| | Teacher engagement | 0.30 | 0.17 | 0.07 | |
| All subjects | Intercept | -1.88 | 0.24 | 0.00 | 0.13 |
| | Implementation (Yr1 vs. Yr2) | 0.69 | 0.32 | 0.03 | 0.23 |
| | Teaching experience | 0.01 | 0.16 | 0.95 | |
| | AP teaching experience | 0.15 | 0.16 | 0.35 | |
| | PSD motivation | -0.11 | 0.16 | 0.51 | |
| | Open enrollment impact | -0.01 | 0.21 | 0.95 | |
| | CRP effectiveness | -0.03 | 0.17 | 0.85 | |
| | Teacher engagement | 0.14 | 0.16 | 0.39 | |

| Outcome | Variable | Est. | S.E | p value | Fitted prob |
|---|---|---|---|---|---|
| Qualifying score | | | | | |
| ELA | Intercept | -6.00 | 0.93 | 0.00 | 0.00 |
| | Implementation (Yr1 vs. Yr2) | -1.88 | 1.27 | 0.14 | |
| | Teaching experience | 0.90 | 0.62 | 0.14 | |
| | AP teaching experience | -0.93 | 0.74 | 0.21 | |
| | PSD motivation | -0.85 | 0.62 | 0.17 | |
| | Open enrollment impact | -0.53 | 0.85 | 0.53 | |
| | CRP effectiveness | 1.46 | 0.67 | 0.03 | 0.01 |
| | Teacher engagement | 0.75 | 0.63 | 0.23 | |
| Math | Intercept | -5.61 | 0.90 | 0.00 | 0.00 |
| | Implementation (Yr1 vs. Yr2) | -0.55 | 1.12 | 0.63 | |
| | Teaching experience | -0.09 | 0.58 | 0.88 | |
| | AP teaching experience | 0.55 | 0.60 | 0.36 | |
| | PSD motivation | -0.21 | 0.58 | 0.72 | |
| | Open enrollment impact | 0.59 | 0.79 | 0.46 | |
| | CRP effectiveness | -0.09 | 0.64 | 0.89 | |
| | Teacher engagement | 0.88 | 0.60 | 0.14 | |
| Science | Intercept | -5.07 | 0.65 | 0.00 | 0.01 |
| | Implementation (Yr1 vs. Yr2) | -0.51 | 0.84 | 0.54 | |
| | Teaching experience | 0.14 | 0.43 | 0.75 | |
| | AP teaching experience | 0.86 | 0.44 | 0.05 | 0.01 |
| | PSD motivation | -0.20 | 0.43 | 0.64 | |
| | Open enrollment impact | 0.98 | 0.59 | 0.10 | |
| | CRP effectiveness | -0.46 | 0.47 | 0.33 | |
| | Teacher engagement | 0.62 | 0.44 | 0.16 | |
| All subjects | Intercept | -4.40 | 0.74 | 0.00 | 0.01 |
| | Implementation (Yr1 vs. Yr2) | -0.62 | 0.94 | 0.51 | |
| | Teaching experience | 0.46 | 0.52 | 0.37 | |
| | AP teaching experience | 0.91 | 0.55 | 0.10 | |
| | PSD motivation | -0.57 | 0.51 | 0.26 | |
| | Open enrollment impact | 0.78 | 0.73 | 0.29 | |
| | CRP effectiveness | 0.11 | 0.55 | 0.84 | |
| | Teacher engagement | 0.64 | 0.54 | 0.24 | |

| Outcome | Variable | Est. | S.E | p value | Fitted prob |
|---|---|---|---|---|---|
| Qualifying score (exams taken) | | | | | |
| ELA | Intercept | -2.00 | 0.63 | 0.00 | 0.12 |
| | Implementation (Yr1 vs. Yr2) | -1.13 | 0.86 | 0.19 | |
| | Teaching experience | 1.21 | 0.46 | 0.01 | 0.31 |
| | AP teaching experience | -0.73 | 0.55 | 0.18 | |
| | PSD motivation | -0.52 | 0.42 | 0.22 | |
| | Open enrollment impact | -0.17 | 0.55 | 0.76 | |
| | CRP effectiveness | 0.83 | 0.44 | 0.06 | |
| | Teacher engagement | 0.45 | 0.44 | 0.30 | |
| Math | Intercept | 0.06 | 0.64 | 0.92 | 0.52 |
| | Implementation (Yr1 vs. Yr2) | -2.31 | 0.86 | 0.01 | 0.09 |
| | Teaching experience | -0.10 | 0.44 | 0.82 | |
| | AP teaching experience | 0.40 | 0.45 | 0.38 | |
| | PSD motivation | -0.50 | 0.40 | 0.21 | |
| | Open enrollment impact | 0.24 | 0.46 | 0.61 | |
| | CRP effectiveness | 0.53 | 0.44 | 0.23 | |
| | Teacher engagement | 1.30 | 0.42 | 0.00 | 0.80 |
| Science | Intercept | -1.12 | 0.68 | 0.10 | 0.25 |
| | Implementation (Yr1 vs. Yr2) | -1.74 | 0.87 | 0.05 | 0.05 |
| | Teaching experience | 0.16 | 0.46 | 0.73 | |
| | AP teaching experience | 0.59 | 0.47 | 0.21 | |
| | PSD motivation | -0.14 | 0.44 | 0.75 | |
| | Open enrollment impact | 0.80 | 0.49 | 0.10 | |
| | CRP effectiveness | -0.06 | 0.44 | 0.88 | |
| | Teacher engagement | 0.42 | 0.42 | 0.32 | |
| All subjects | Intercept | -1.55 | 0.73 | 0.03 | 0.18 |
| | Implementation (Yr1 vs. Yr2) | -1.41 | 0.94 | 0.13 | |
| | Teaching experience | 0.58 | 0.49 | 0.24 | |
| | AP teaching experience | 0.70 | 0.51 | 0.17 | |
| | PSD motivation | -0.44 | 0.46 | 0.34 | |
| | Open enrollment impact | 0.68 | 0.55 | 0.22 | |
| | CRP effectiveness | 0.15 | 0.46 | 0.75 | |
| | Teacher engagement | 0.51 | 0.44 | 0.25 | |

Table 10 provides the HGLM results of the relationships between the three outcome measures and the implementation covariates using 2019 AP exam data. As with the 2017–2018 results, for each model presented, the fitted probability of the intercept corresponds with the likelihood of the outcome for the Delayed Treatment Schools. There were no statistically significant differences in any of our 2018–2019 outcome measures between the two groups of schools, holding the other variables in the model constant.

There were no significant results for the first outcome, "Total exams," across all subjects, concerning relationships between covariates and the probability of taking at least one AP exam. For the math subject category, a unit increase in "AP teaching experience" was associated with a 3% increase in the fitted probability of taking at least one Math AP exam (Est. = 0.56, S.E. = 0.26, fitted probability = 0.08). For science AP exams, "PSD motivation" was negatively related to the outcome, with a unit increase related to a 3% decrease in the fitted probability (Est. = -0.58, S.E. = 0.25, fitted probability = 0.03).

The second outcome, "Qualifying score" was significantly and positively associated with "AP teaching experience" for all subjects (Est. = 1.44, S.E. = 0.53, fitted probability = 0.05). "AP teaching experience" was also a significant predictor for the ELA and math subject domains. For ELA, "AP teaching experience" only had a marginal difference in fitted probability associated with a unit change in the covariate (Est. =1.66, S.E. = 0.76, fitted probability = 0.00). The result is similar to Math (Est. =1.46, S.E. = 0.58, fitted probability = 0.01).

For the last outcome, "Qualifying score (exams taken)," a significant and positive association between "AP teaching experience" and the outcome was found across all subjects (Est. = 1.25, S.E. = 0.48, fitted probability = 0.29), as well as for ELA and math AP exams. In ELA, a unit increase in "AP teaching experience" was related to a 14% increase in the fitted probability of the outcome (Est. =1.75, S.E. = 0.76, fitted probability = 0.18). In math, "AP teaching experience" was related to a 28% increase in the fitted probability to the outcome (Est. = 1.46, S.E. = 0.47, fitted probability = 0.43). In addition, a unit increase in "Open enrollment impact" was related to a 10% increase in the fitted probability (Est. = 0.64, S.E. = 0.30, fitted probability difference = 0.25).

Table 10

*2-Level HGLM Results: 2019 AP Exam Data*

| Outcome | Variable | Est. | S.E | p value | Fitted prob |
|---|---|---|---|---|---|
| Total exams | | | | | |
| ELA | Intercept | -3.56 | 0.37 | 0.00 | 0.03 |
| | Implementation (Yr1 vs. Yr2) | 0.29 | 0.55 | 0.59 | |
| | Teaching experience | -0.26 | 0.27 | 0.34 | |
| | AP teaching experience | 0.13 | 0.26 | 0.62 | |
| | PSD motivation | 0.27 | 0.31 | 0.37 | |
| | Open enrollment impact | -0.17 | 0.30 | 0.56 | |
| | CRP effectiveness | -0.49 | 0.31 | 0.11 | |
| | Teacher engagement | -0.11 | 0.23 | 0.64 | |
| Math | Intercept | -3.03 | 0.38 | 0.00 | 0.05 |
| | Implementation (Yr1 vs. Yr2) | -0.96 | 0.59 | 0.10 | |
| | Teaching experience | -0.17 | 0.28 | 0.55 | |
| | AP teaching experience | 0.56 | 0.26 | 0.03 | 0.08 |
| | PSD motivation | 0.01 | 0.32 | 0.97 | |
| | Open enrollment impact | -0.14 | 0.32 | 0.66 | |
| | CRP effectiveness | 0.17 | 0.32 | 0.59 | |
| | Teacher engagement | -0.46 | 0.25 | 0.07 | |
| Science | Intercept | -2.75 | 0.31 | 0.00 | 0.06 |
| | Implementation (Yr1 vs. Yr2) | 0.34 | 0.46 | 0.46 | |
| | Teaching experience | 0.00 | 0.21 | 1.00 | |
| | AP teaching experience | 0.14 | 0.19 | 0.46 | |
| | PSD motivation | -0.58 | 0.25 | 0.02 | 0.03 |
| | Open enrollment impact | -0.04 | 0.24 | 0.86 | |
| | CRP effectiveness | 0.21 | 0.24 | 0.37 | |
| | Teacher engagement | -0.10 | 0.18 | 0.59 | |
| All subjects | Intercept | -1.71 | 0.28 | 0.00 | 0.15 |
| | Implementation (Yr1 vs. Yr2) | 0.07 | 0.42 | 0.86 | |
| | Teaching experience | -0.11 | 0.18 | 0.53 | |
| | AP teaching experience | 0.23 | 0.17 | 0.18 | |
| | PSD motivation | -0.21 | 0.21 | 0.31 | |
| | Open enrollment impact | 0.00 | 0.21 | 1.00 | |
| | CRP effectiveness | 0.09 | 0.21 | 0.67 | |
| | Teacher engagement | -0.13 | 0.16 | 0.41 | |

| Outcome | Variable | Est. | S.E | p value | Fitted prob |
|---|---|---|---|---|---|
| Qualifying score | | | | | |
| ELA | Intercept | -7.04 | 0.95 | 0.00 | 0.00 |
| | Implementation (Yr1 vs. Yr2) | 0.59 | 1.25 | 0.64 | |
| | Teaching experience | -0.10 | 0.74 | 0.90 | |
| | AP teaching experience | 1.66 | 0.76 | 0.03 | 0.00 |
| | PSD motivation | 0.43 | 0.90 | 0.63 | |
| | Open enrollment impact | 0.37 | 0.74 | 0.61 | |
| | CRP effectiveness | -1.33 | 0.90 | 0.14 | |
| | Teacher engagement | -0.30 | 0.64 | 0.64 | |
| Math | Intercept | -6.09 | 0.80 | 0.00 | 0.00 |
| | Implementation (Yr1 vs. Yr2) | -0.07 | 1.07 | 0.95 | |
| | Teaching experience | 0.01 | 0.57 | 0.98 | |
| | AP teaching experience | 1.87 | 0.58 | 0.00 | 0.01 |
| | PSD motivation | -0.20 | 0.68 | 0.77 | |
| | Open enrollment impact | 0.91 | 0.62 | 0.14 | |
| | CRP effectiveness | -0.25 | 0.67 | 0.71 | |
| | Teacher engagement | -0.16 | 0.50 | 0.75 | |
| Science | Intercept | -5.91 | 0.92 | 0.00 | 0.00 |
| | Implementation (Yr1 vs. Yr2) | 0.27 | 1.25 | 0.83 | |
| | Teaching experience | -0.94 | 0.72 | 0.19 | |
| | AP teaching experience | 0.98 | 0.64 | 0.12 | |
| | PSD motivation | -0.68 | 0.71 | 0.34 | |
| | Open enrollment impact | -0.01 | 0.74 | 0.99 | |
| | CRP effectiveness | -0.41 | 0.69 | 0.56 | |
| | Teacher engagement | -0.25 | 0.55 | 0.65 | |
| All subjects | Intercept | -4.48 | 0.73 | 0.00 | 0.01 |
| | Implementation (Yr1 vs. Yr2) | 0.05 | 1.05 | 0.96 | |
| | Teaching experience | -0.60 | 0.55 | 0.27 | |
| | AP teaching experience | 1.44 | 0.53 | 0.01 | 0.05 |
| | PSD motivation | -0.80 | 0.58 | 0.16 | |
| | Open enrollment impact | -0.34 | 0.56 | 0.55 | |
| | CRP effectiveness | 0.11 | 0.58 | 0.86 | |
| | Teacher engagement | -0.34 | 0.45 | 0.45 | |

| Outcome | Variable | Est. | S.E | p value | Fitted prob |
|---|---|---|---|---|---|
| Qualifying score (exams taken) | | | | | |
| ELA | Intercept | -3.24 | 0.97 | 0.00 | 0.04 |
| | Implementation (Yr1 vs. Yr2) | 1.06 | 1.26 | 0.40 | |
| | Teaching experience | 0.22 | 0.70 | 0.75 | |
| | AP teaching experience | 1.75 | 0.76 | 0.02 | 0.18 |
| | PSD motivation | -0.02 | 0.75 | 0.98 | |
| | Open enrollment impact | 0.56 | 0.62 | 0.37 | |
| | CRP effectiveness | -0.59 | 0.70 | 0.40 | |
| | Teacher engagement | -0.13 | 0.60 | 0.83 | |
| Math | Intercept | -1.76 | 0.59 | 0.00 | 0.15 |
| | Implementation (Yr1 vs. Yr2) | 1.18 | 0.87 | 0.18 | |
| | Teaching experience | 0.31 | 0.48 | 0.52 | |
| | AP teaching experience | 1.46 | 0.47 | 0.00 | 0.43 |
| | PSD motivation | -0.63 | 0.51 | 0.21 | |
| | Open enrollment impact | 0.64 | 0.30 | 0.03 | 0.25 |
| | CRP effectiveness | -0.15 | 0.46 | 0.74 | |
| | Teacher engagement | 0.28 | 0.39 | 0.46 | |
| Science | Intercept | -2.67 | 0.88 | 0.00 | 0.06 |
| | Implementation (Yr1 vs. Yr2) | -0.13 | 1.25 | 0.92 | |
| | Teaching experience | -1.00 | 0.71 | 0.16 | |
| | AP teaching experience | 0.92 | 0.65 | 0.15 | |
| | PSD motivation | -0.16 | 0.63 | 0.79 | |
| | Open enrollment impact | 0.03 | 0.53 | 0.95 | |
| | CRP effectiveness | -0.68 | 0.65 | 0.29 | |
| | Teacher engagement | -0.27 | 0.53 | 0.62 | |
| All subjects | Intercept | -2.14 | 0.65 | 0.00 | 0.11 |
| | Implementation (Yr1 vs. Yr2) | 0.14 | 0.95 | 0.89 | |
| | Teaching experience | -0.41 | 0.51 | 0.42 | |
| | AP teaching experience | 1.25 | 0.48 | 0.01 | 0.29 |
| | PSD motivation | -0.65 | 0.47 | 0.16 | |
| | Open enrollment impact | -0.21 | 0.37 | 0.57 | |
| | CRP effectiveness | -0.04 | 0.46 | 0.93 | |
| | Teacher engagement | -0.31 | 0.40 | 0.44 | |

## V. Results for Implementation Evaluation

The fidelity of implementation matrix defines the key components of the CRP program depicted in the CRP logic model (see Figure 1), measures of each component, scoring rubrics of measures, and criteria of fidelity. Based on this fidelity matrix, implementation information was collected from administrative records, surveys, and interviews.

The fidelity matrix approach collected information based on observable and measurable indicators relating to key program features. The CRP logic model posits that the key components of the intervention are school, teacher, and student supports. The idea was to measure fidelity separately for each key component of the intervention and define threshold values (in collaboration with NMSI) to determine whether the intervention was implemented with fidelity. As a starting point, we used fidelity indicators developed and field-tested (e.g., Sherman et al., 2015), and created an implementation fidelity matrix which linked the key components of the intervention to their indicators, the data source, the indicator scoring system, and the implementation threshold values.

Fidelity was measured separately for each key component of the intervention, and the components scores summed to determine whether the intervention was implemented overall with fidelity. The key components were further categorized as either involving NMSI fulfilling administrative requirements or measuring the participation of students or school personnel. The latter distinction provided a more detailed and formative view of the data. We realized that scores could not be interpreted in a meaningful way if the administrative components were added to the school and teacher participation components. For example, of the eight "teacher" matrix elements, four assess activities that directly influence classroom instruction and the remaining four evaluate NMSI's fulfillment of administrative responsibilities. If a school received 6/8 "teacher-related points" we would now know if the majority of the points came from activities directly influencing classroom instruction, vs. those relating to NSMI's contributions.

In some cases implementation could be measured on a yes/no basis (e.g., did schools receive necessary materials, or were AP exam fees paid by NMSI?). These elements were considered to have been implemented with fidelity if in at least 80% of the schools they were implemented as planned. Some program elements required a number of individuals participating in an event (e.g., attending a summer training session, or attending three student study sessions). In these cases, if 80% of identified staff or students attended, these elements were considered to have been implemented with fidelity.

We also collected implementation information from surveys and interviews. Interview and survey data collected for the study allow us to determine how components of the program

function in "real-world contexts," and provide support and validation for the fidelity matrix data. More importantly, these data help us learn how components of the program are viewed by those within the school, and how teachers can be supported so programs function effectively and lead to positive change. Obtaining teachers' firsthand views and opinions provides unique insight into what is necessary to build and sustain an effective AP program and supportive school culture, and if need be, provide formative feedback to help guide modifications or enhancements. Teachers are more likely to understand the complexity of their particular school, classroom, and student population and are most closely connected to actual program implementation.

To determine the perceived effectiveness of program elements we created an online survey and interview protocol. The survey focused on the key components of the CRP: teacher training, student supports, administrative support, monetary incentives, additional instruction (via student study sessions), classroom supplies/equipment, and change in school culture as relates to the AP program. The online survey questions were used as a basis for the interview protocol and allowed us to expand on topics in the survey. The measure was based on one created for previous studies of the CRP (Cross et al., 2012), with additional items added for this project.

The online surveys were created using the web-based survey creation program Survey Monkey. Alternate versions were created for teachers and two categories of school administrator (based on roles associated with the CRP). Prior to sending the surveys, we sent an introductory email explaining the research study, participation requirements, as well as the time frame for participation. Eligible participants were subsequently emailed a link to the survey with the study overview and an IRB-required consent form. Prior to the response deadline, several reminder emails were sent.

## A. 2017–2018 Fidelity Matrix Data Collection

### a. 2017–2018 Teacher and Administrator Surveys

The majority of survey questions were multiple choice, multiple select (respondents could choose more than one answer), or questions with a Likert scale (most often 4-point). Skip logic was employed to allow participants to move past a set of questions if they were not relevant or applicable.

The 2018 teacher survey contained 72 questions, although all participants did not answer every question. The first section of the survey included demographic questions (respondent's school, courses taught, grade level etc.) as well as questions on professional development participation. The rest of the survey sections were aligned to the CRP logic model constructs and addressed school support, teacher support, and student support.

The Partner School Director (the role formerly known as the Designated Administrator) survey was similarly constructed, albeit aligned to the administrative elements of the logic model. Thus, there were questions focused on coordination of the CRP, student supports, and administrator incentives, as well as identical questions to those on the teacher survey about effectiveness of program components on enrollment in AP courses and performance on AP exams.

The Site Coordinator survey primarily focused on the coordinator's role in organizing student study sessions. In addition, we gathered information about the effectiveness of program components.

### b. 2017–2018 Student Survey

Unlike the online teacher and administrator surveys, the student surveys were paper booklets designed to be completed by hand in class. Following completion of the online teacher survey, a packet of student surveys specifically coded for the teacher and school were sent to the teacher with a postage-paid return envelope. The survey consisted of 19 questions about the student's experience with the CRP and AP courses in general.

The survey began with five questions about the student's current AP workload, college aspirations, and the school's outreach efforts to students concerning AP course enrollment. The students were then asked eight questions about the supports available to them, not limited to CRP components, including two questions intended to solicit qualitative assessments of the supports.

The survey continued with two questions about the student's perception of their own preparedness and potential obstacles to students taking their AP exams. The final four questions identified the various incentives offered to students and gauged the student's valuation of them. The incentives probed included CRP financial awards and other perceived benefits, such as weighted calculations of grade point average.

### c. 2017–2018 Teacher and Administrator Interviews

As a supplement to the surveys we created a set of teacher and administrator interview questions to provide more detailed information which may not have been gleaned from the online survey. The scripts formed the basis of interviews with teachers and administrators, and additional questions were included as dictated by the nature of the conversation. We did not want to constrain teachers and so allowed the discussions to evolve as they progressed.

## B. 2017–2018 Fidelity Matrix Results

The fidelity matrix relates to all components of the CRP, includes targets for each element of the program and requires 80% of schools to meet specified targets. Incomplete data made

completing the matrix in the way intended challenging. We did not receive Partner School Director surveys from 15 schools, student surveys from 14 schools, Site Coordinator surveys from eight schools, and had incomplete student attendance records for the student study sessions (see Table 11). Data to complete the matrix were gathered from administrative records as well as survey and interview responses, where necessary.

Table 11

*2017–2018 Survey and Interview Participation*

| Instrument | Number completed | Schools represented | % of eligible schools |
|---|---|---|---|
| Teacher Survey | 200 | 48 | 100% |
| Partner School Director Survey | 33 | 33 | 69% |
| Site Coordinator Survey | 44 | 40 | 83% |
| Student Survey | 1,930 | 34 | 71% |
| Teacher Interview | 85 | 24 | 50% |
| Partner School Director Interview | 14 | 14 | 29% |

### a. School Indicators: 2017–2018 Fidelity Matrix

School indicators included program management support, school-wide goal setting, and payment of administrator awards. Each of the seven school key metrics were evaluated on a yes/no basis and measured NMSI's proficiency in meeting administrative responsibilities relating to the CRP.

School personnel in 43/48 Program Schools (89.6%) confirmed that a goal setting meeting took place. Seventy percent of Partner School Directors reported that teachers at their school participated in establishing the CRP goals for their classes, which contradicts teacher interviews during which 88% of teachers said they *were not* consulted about establishing the goals.

NMSI financial records indicated that 89.6% of the Site Coordinator stipends and 100% of the Partner School Director bonuses were paid to all qualifying schools. Forty-one schools indicated that NMSI Program Managers fulfilled all four expected functions (listed in Table 12), five schools indicated support in three categories, and two schools reported support in only two categories. Table 12 summarizes the school support measures across all schools and reveals that each program element was implemented with fidelity.

Table 12

*School Implementation Indicators*

| School components | # of schools | # compliant | % compliant |
|---|---|---|---|
| Goal Meeting Held | 48 | 43 | 89.6 |
| Site Coordinator Stipend Paid | 48 | 43 | 89.6 |
| PSD Bonus Paid According to Agreement | 39 | 39 | 100.0 |
| Program Manager Support: Teacher Sign-Ups | 48 | 47 | 97.9 |
| Program Manager Support: Student Study Sessions | 48 | 47 | 97.9 |
| Program Manager Support: NMSI Materials | 48 | 47 | 97.9 |
| Program Manager Support: Mentor Assignment | 48 | 42 | 87.5 |

### b. Teacher Indicators: 2017–2018 Fidelity Matrix

Seven specific metrics were identified to evaluate (a) teacher participation in the CRP (one component) and (b) NMSI support of teachers (six components). Nine schools (18.8%) achieved a perfect score on the teacher support program elements. Table 13 summarizes the range of scores in the teacher measures.

Table 13

*Measures of Teacher Participation and Support*

| Category | Score of 5 | Score of 6 | Score of 7 |
|---|---|---|---|
| # of schools | 7 | 32 | 9 |
| % of schools | 14.6% | 66.7% | 18.8% |

The CRP supports teacher instruction by providing materials for the classroom, online access to instructional resources, mentoring, and off-site training. Financial records indicated that each school received funding to purchase school room materials and supplies, ranging from books to lab supplies. Teachers in 45 of the schools (93.8%) responded affirmatively when asked if they had the materials they need to teach their course effectively. Teachers in all 48 schools were given access to online instructional resources and logged into the system at least once. Mentors were also offered to teachers in 44 schools.

Schools were evaluated on the level of teacher participation in CRP training sessions by assessing whether teachers attended all three sessions available throughout the year (Advanced Placement Summer Institute, Fall Workshop, and Spring Training). If 80% or above of teachers

in a given school attended all three training sessions, the school received a score of one and this was achieved by 11 schools (22.9%). Some teachers cited the weeklong commitment as an obstacle to attending the summer training, whereas other roles and responsibilities that teachers maintain (e.g., coaching, advising student organizations) can also make attendance during the school year challenging. Only four schools (8.3%) had 100% teacher attendance at training sessions.

The CRP also offers financial support to teachers for participating in the program, both through a stipend and through awards tied to achieving goals. NMSI financial records confirmed that teachers at all 48 schools received stipend payments for program participation. For each student with an exam score of 3 or higher, teachers should have received an award of $100. Payments to teachers according to the agreement were verified for 47/48 schools (97.9%).

In many schools ($n$ = 37) teachers did not fulfill their requirements for attending all required training sessions, and so this component was not implemented with fidelity. The other six program elements were implemented with fidelity in the aggregate across all schools.

### c. Student Supports: 2017–2018 Fidelity Matrix

Among the student supports in the Fidelity Matrix are three financial components: purchasing classroom supplies or materials, subsidizing AP exam fees, and rewarding students for qualifying scores. All 48 schools received funds to purchase supplies and materials for classrooms and labs. By agreement, not all districts were expected to receive exam fee subsidies because in some districts the exams are subsidized by other sources. The 14 schools in these districts were considered compliant, as were 31 other schools for which the subsidies were paid as expected—for a total of 45 compliant schools (93.8%). Students who achieve a qualifying score on the AP exam are eligible for a $100 award from NMSI. Students in all 48 schools received the award commensurate with performance.

Students are expected to attend three student study sessions for each of the CRP AP courses in which they are enrolled. Student study sessions are subject-specific opportunities on Saturday mornings for all students in participating schools in a region to receive instruction together from a visiting, experienced AP teacher. However, on occasion, sessions are canceled and may not be rescheduled and attendance tracking is inconsistent. In the student survey, participants were asked how many student study sessions were offered for the course, and how many of the sessions they attended. For this metric, compliance was determined at the school level based on the number of students who self-reported attending all of the available sessions for the course. If 80% or more of the students surveyed attended all of the sessions, the school was in compliance. In the teacher and site coordinator surveys, participants were asked what

percentage of students attended three student study sessions per course. If a survey respondent selected “75%-100%” the school was considered in compliance. Twenty-four schools met one or both of these two criteria.

Four total points could be earned from the student components, and 22 schools earned all four (45.8%). See Table 14.

Table 14

*Measures of Student Supports*

| Category | Score of 2 | Score of 3 | Score of 4 |
|---|---|---|---|
| # of schools | 1 | 25 | 22 |
| % of schools | 2.1% | 52.1% | 45.8% |

### d. Overall School-Level Fidelity: 2017–2018 Fidelity Matrix

Each of the measures was tallied per school, and the school’s total score evaluated as a percentage of the maximum points available (18). Forty-three schools (89.6%) achieved 80% or better implementation fidelity, for an average fidelity score of 89.4%. Four schools achieved a perfect 100% fidelity score (see Table 15).

Table 15

*School-Level Fidelity Matrix Scores*

| Overall total | Number of schools |
|---|---|
| Schools with a score of 18 | 4 |
| Schools with a score of 17 | 17 |
| Schools with a score of 16 | 16 |
| Schools with a score of 15 | 6 |
| Schools with a score of 14 | 2 |
| Schools with a score of 13 | 2 |
| Schools with a score of 12 | 0 |
| Schools with a score of 11 | 1 |

In the aggregate by component, the percentage of schools that adequately implemented the measures ranged from 45.8% for the student measures to 93.8% for the school-level measures.

Table 16 delineates by component the number and percentage of schools performing at or above 80% fidelity.

Table 16

*Schools Adequately Implementing Aggregate Measures by Component*

| Component | # of schools | # adequately implemented (at or above 80%) | % adequately implemented |
|---|---|---|---|
| School-level indicators | 48 | 45 | 93.8 |
| Teacher-level indicators | 48 | 41 | 85.4 |
| Student-level indicators | 48 | 22 | 45.8 |

## C. 2017–2018 Survey Results

The research objectives for this part of the evaluation study were to better understand teacher and administrator perspectives on the effectiveness or impact of key CRP components on student interest and success in AP and on school culture. For all survey questions, frequencies for each response category were first calculated. Some questions required respondents to respond using Likert (4 or 3 points) or dichotomous scales. For these questions we calculated mean values to gain an understanding of average level of agreement with the statements. Some questions shared a common prompt. For these questions, we further computed the average of the set of item responses as a composite score per respondent and obtained descriptive statistics.

### a. Response Rates: 2017–2018 Surveys

Completed surveys were received from 200 teachers (around 62% of the eligible sample). We received a teacher survey from at least one teacher in each of the 48 schools in the sample, with a 100% response rate from 10 schools. Of the 200 respondents 110 were in their first year of CRP implementation, and 90 were in their second year. Partner School Director surveys were received from 33 administrators. Table 17 provides the distributions of teachers responding to the survey by state, district, and school.

Table 17

*Distribution of Teacher Survey Respondents by State, District, and School*

| State | District | School | *n* | % |
|---|---|---|---|---|
| CA | CA district | California 2 | 3 | 1.5 |
| | | California 3 | 7 | 3.5 |
| GA | GA district | Georgia 1 | 1 | 0.5 |
| | | Georgia 2 | 7 | 3.5 |
| | | Georgia 3 | 4 | 2.0 |
| | | Georgia 4 | 6 | 3.0 |
| IL | IL district | Illinois 1 | 5 | 2.5 |
| | | Illinois 2 | 6 | 3.0 |
| | | Illinois 3 | 7 | 3.5 |
| | | Illinois 4 | 1 | 0.5 |
| | | Illinois 5 | 4 | 2.0 |
| LA | LA district | Louisiana 1 | 9 | 4.5 |
| | | Louisiana 2 | 7 | 3.5 |
| MI | MI district | Michigan 1 | 7 | 3.5 |
| | | Michigan 2 | 6 | 3.0 |
| | | Michigan 3 | 1 | 0.5 |
| MO | MO district | Missouri 1 | 4 | 2.0 |
| | | Missouri 2 | 2 | 1.0 |
| | | Missouri 3 | 4 | 2.0 |
| | | Missouri 4 | 2 | 1.0 |
| | | Missouri 5 | 6 | 3.0 |
| | | Missouri 6 | 1 | 0.5 |
| ND | ND district 1 | North Dakota 1 | 5 | 2.5 |
| | | North Dakota 2 | 8 | 4.0 |
| | | North Dakota 3 | 5 | 2.5 |
| | ND district 2 | North Dakota 4 | 5 | 2.5 |
| | | North Dakota 5 | 7 | 3.5 |

| State | District | School | *n* | % |
|---|---|---|---|---|
| OH | OH district | Ohio 1 | 1 | 0.5 |
| | | Ohio 2 | 2 | 1.0 |
| | | Ohio 3 | 3 | 1.5 |
| | | Ohio 4 | 5 | 2.5 |
| | | Ohio 5 | 3 | 1.5 |
| | | Ohio 6 | 3 | 1.5 |
| | | Ohio 7 | 1 | 0.5 |
| | | Ohio 8 | 2 | 1.0 |
| PA | PA district 1 | Pennsylvania 1 | 3 | 1.5 |
| | PA district 2 | Pennsylvania 2 | 8 | 4.0 |
| | PA district 3 | Pennsylvania 3 | 5 | 2.5 |
| | PA district 4 | Pennsylvania 4 | 9 | 4.5 |
| | PA district 5 | Pennsylvania 5 | 8 | 4.0 |
| TX | TX district | Texas 1 | 1 | 0.5 |
| | | Texas 2 | 1 | 0.5 |
| | | Texas 3 | 3 | 1.5 |
| | | Texas 4 | 1 | 0.5 |
| | | Texas 5 | 3 | 1.5 |
| | | Texas 6 | 3 | 1.5 |
| | | Texas 7 | 3 | 1.5 |
| | | Texas 8 | 2 | 1.0 |

### b. 2017–2018 Teacher Survey Response Summary

Over 95% of teacher respondents had three or more years of teaching experience. Sixty-nine teachers had a single subject credential and 94 had multiple subject credentials. The majority of respondents taught 12th grade (*n* = 175) and/or 11th grade (*n* = 169). Ninety-eight teachers taught 10th grade and 61 taught ninth grade. On average, participants had been teaching for 13.92 years (*SD* = 8.78, *Mdn* = 13.00), and had been teaching AP courses an average of 4.77 years (*SD* = 4.93, *Mdn* = 3.00).

#### 1. Training: 2017–2018 Teacher Survey

We asked teachers which of a set of statements was true for each of the three training sessions they attended. Specifically we were interested in how effective teachers found the training sessions for achieving stated goals. Table 18 depicts the results of responses to these

survey items. In all cases, fewer teachers found the training sessions effective in helping them differentiate instruction for students at different ability levels (between 48% and 58%).

Table 18

*Teacher Evaluation of CRP Training Activities*

| Training session evaluation | Summer Session | | Fall Session | | Spring Session | |
|---|---|---|---|---|---|---|
| | # | % | # | % | # | % |
| Attended | 176 | | 184 | | 174 | |
| Scheduling & Location Convenient | 108 | 61 | 118 | 64 | 119 | 68 |
| Knowledgeable & Well-Prepared Facilitators | 155 | 88 | 152 | 83 | 150 | 86 |
| Improved My Content Knowledge | 138 | 78 | 130 | 71 | 112 | 64 |
| Felt More Qualified | 135 | 77 | 130 | 71 | 119 | 68 |
| Clear Agenda & Goals | 152 | 86 | 149 | 81 | 140 | 80 |
| Effective Training Activities | 136 | 77 | 132 | 72 | 129 | 74 |
| Helped Me Differentiate Instruction | 103 | 59 | 88 | 48 | 84 | 48 |

More generally, nearly half of the survey participants felt that the four-day summer institute was the most beneficial component of the College Readiness Program (see Table 19).

Table 19

*Most Beneficial Component of the CRP*

| Most beneficial program element (*n*=194) | Frequency | Percent |
|---|---|---|
| CRP Summer Institute | 96 | 49% |
| Student Study Sessions | 34 | 17% |
| Fall Training | 29 | 14% |
| Spring Training | 27 | 13% |
| Mentoring | 8 | 4% |

A higher percentage of teachers in their second year of the program felt that summer training helped them differentiate instruction (62%) compared to teachers in the first year of the program (54%). Similarly a higher percentage of second year teachers believed that the fall training helped them differentiate instruction (52%) and improved their content knowledge (76%) compared to year one teachers (45% and 68%, respectively). For the spring training, a

higher percentage of second year teachers felt positively about every aspect of the training than did first year teachers (see Figure 5).

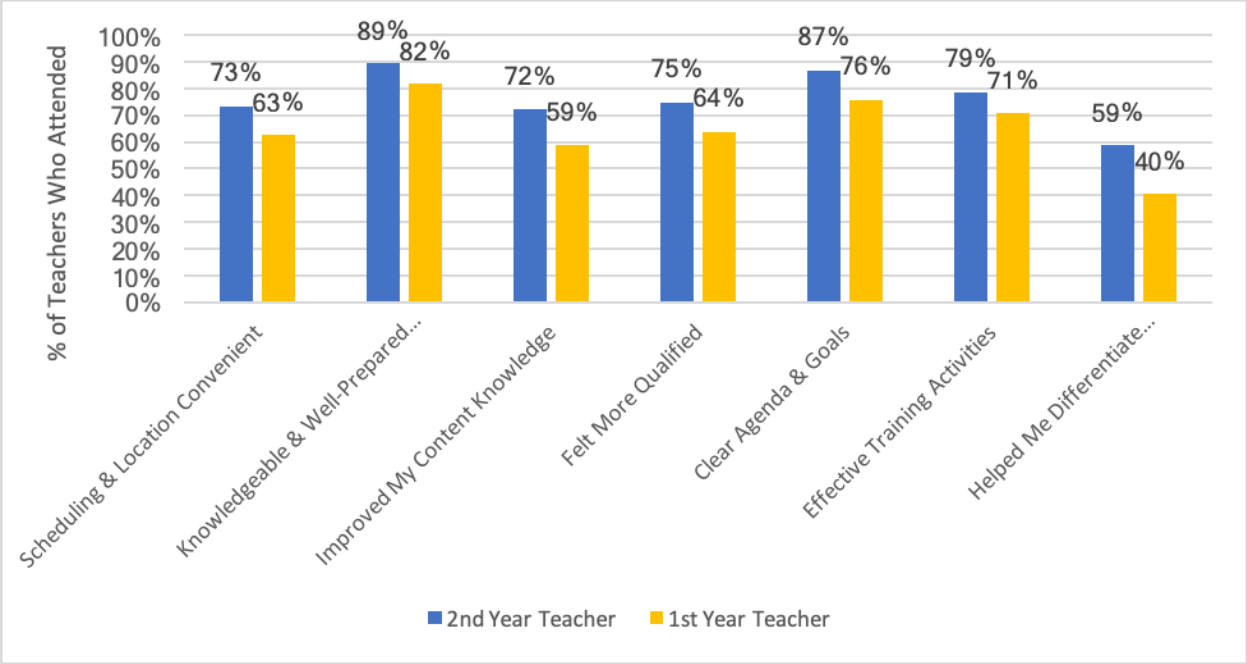



Figure 5. Teacher evaluations of spring training by teachers' years in the CRP.

We further asked second year teachers how the training sessions had changed from the 2016–2017 school year. While nearly a quarter of respondents either did not know if the training had changed or felt that the training had not changed (7% and 17%, respectively), many of the other participants noted that the content had changed, teacher participation had increased, and the quality of instruction had improved (see Table 20).

Table 20

*Changes to Teacher Training From SY2016–2017*

| Ways in which teacher training changed from 2016–2017 school year | Frequency | Percent |
|---|---|---|
| Addressed new/different topics | 34 | 26% |
| No changes | 23 | 17% |
| Increased teacher participation | 18 | 13% |
| Quality of instructors has improved | 17 | 13% |
| Increased utility | 12 | 9% |
| Scheduling has improved | 11 | 8% |
| Do not know | 10 | 7% |
| Other | 5 | 3% |

*Note.* Limited to second year teachers (multiple choice, multiple selection).

We also asked teachers to indicate how much the CRP had improved their content knowledge, instructional skills, techniques and strategies. More than half of the respondents (56%) felt that the CRP contributed to a major improvement in content knowledge (average rating was 2.50 on a 3-point scale: 3 = *major improvement*, 2 = *slight improvement*, 1 = *no improvement*). Similarly, 60% of teachers indicated improvement in their instructional skills and strategies following the CRP (*M* = 2.54).

The student study sessions offered opportunities for teachers to observe expert teachers and learn new instructional techniques or strategies. We asked teachers to rate the extent of their agreement on a four-point scale with a set of statements about the student study sessions. The mean of teachers' responses ranged from 2.95 to 3.22 (see Table 21).

Table 21

*Teacher Ratings of Student Study Session Efficacy*

| Student study session statements | *n* | *M* | *SD* |
|---|---|---|---|
| The study sessions highlighted the instructional needs of the students so I could continue to address them in class. | 199 | 3.22 | 0.76 |
| I was able to take the strategies employed during the study sessions back to the classroom to help improve student achievement. | 197 | 3.07 | 0.90 |
| I learned a great deal from watching the expert teachers. | 197 | 2.95 | 0.95 |

*Note*. 1 = *Strongly Disagree*, 2 = *Disagree Somewhat*, 3 = *Agree Somewhat*, 4 = *Strongly Agree*.

A separate item asked teachers to indicate how useful the sessions were for them (see Table 22).

Table 22

*Teacher Ratings of Student Study Session Usefulness*

| Student study session statements | *n* | Mean | SD |
|---|---|---|---|
| How useful were study sessions for you? | 196 | 2.90 | 0.98 |

*Note.* 1 = *not at all useful*, 2 = *somewhat useful*, 3 = *useful*, 4 = *extremely useful*.

More than two thirds of teachers (68%) indicated that the student study sessions were useful or extremely useful. Similarly, teachers expressed consistent views on the efficacy of the specific components of the study sessions: over 71% of teachers reported somewhat or strong agreement.

We tracked individual teacher responses from year to year: in 2018, 83 teachers who had completed the survey in 2017, completed the survey again. When we compared these participants' opinions about student study sessions from the 2016–2017 school year to those of the 2017–2018 school year, we see that ratings improved significantly in three measures: quality of instructors, improving student content knowledge, and highlighting instructional needs of students. The only category that did not show improvement, significant or otherwise, year over year was considering teacher input when determining topics to cover at the student study session (see Table 23).

Table 23

*Changes in Teacher Opinions About Student Study Sessions From First to Second Year in the CRP*

| **Statements about the student study sessions** | **2017** ***M (SD)*** | **2018** ***M (SD)*** | **2017–2018 diff** | **S.E.** | ***t*** | **df** | ***p*** |
|---|---|---|---|---|---|---|---|
| The study sessions were led by AP experts who taught NMSI-created lessons | 3.22 (1.03) | 3.59 (0.53) | -0.373 | 0.146 | -2.504 | 59 | 0.015 |
| The study sessions helped to increase student confidence | 3.13 (0.91) | 3.34 (0.68) | -0.211 | 0.134 | -1.492 | 59 | 0.141 |
| The study sessions improved students' content knowledge | 3.1 (0.95) | 3.4 (0.61) | -0.293 | 0.139 | -2.042 | 59 | 0.046 |
| Students were active participants (e.g., answering and asking questions, focused on tasks assigned, etc.) | 3.034 (0.91) | 3.23 (0.69) | -0.196 | 0.117 | -1.592 | 58 | 0.117 |
| I was able to take the strategies I saw employed during the study sessions back to my own classrooms to improve student achievement | 2.966 (0.96) | 3.1 (0.96) | -0.134 | 0.139 | -1.119 | 57 | 0.268 |
| The study sessions highlighted the instructional needs of the students for me to continue addressing in class | 2.93 (0.94) | 3.23 (0.62) | -0.3 | 0.137 | -2.105 | 58 | 0.04 |
| I learned a great deal from watching the expert teachers during the student study sessions | 2.82 (0.96) | 2.92 (0.89) | -0.1 | 0.142 | -0.599 | 58 | 0.552 |
| The study sessions were conveniently scheduled to accommodate student schedules | 2.85 (0.92) | 3 (0.75) | -0.15 | 0.129 | -1.033 | 59 | 0.306 |
| My input was considered when determining the study session topics | 2 (1.174) | 1.8 (1.04) | 0.2 | 0.192 | 1.166 | 57 | 0.248 |

*Note.* 1 = *Strongly Disagree*, 2 = *Disagree Somewhat*, 3 = *Agree Somewhat*, 4 = *Strongly Agree*.

We also looked for differences in opinion about the student study sessions between the teachers in the Treatment Schools and those in the Delayed Treatment Schools. Across all measures, survey participants from Treatment Schools felt more positively about the student study sessions, with several significant differences (see Table 24). Although the overall trends between groups on the different items were the same. For example, both groups had less agreement on their input being considered for determining session topics, and higher agreement that sessions improved students' content knowledge.

Table 24

*Treatment School and Delayed Treatment School Opinions About Student Study Sessions*

| Statements about the student study sessions | Delayed Treatment *M (SD)* | Treatment *M (SD)* | DT - T diff | S.E. | *t* | Df | *p* |
|---|---|---|---|---|---|---|---|
| My input was considered when determining the study session topics | 1.74 (1.1) | 1.78 (0.97) | -0.044 | 0.15 | -0.289 | 137.97 | 0.773 |
| Students were active participants (e.g., answering and asking questions, focused on tasks assigned, etc.) | 3.01 (0.87) | 3.26 (0.68) | -0.246 | 0.12 | -2.054 | 118.44 | **0.042** |
| The study sessions were conveniently scheduled to accommodate student schedules | 2.69 (0.90) | 3.02 (0.82) | -0.326 | 0.128 | -2.542 | 135.61 | **0.012** |
| The study sessions improved students' content knowledge | 3.24 (0.76) | 3.37 (0.65) | -0.137 | 0.107 | -1.282 | 130.56 | 0.202 |
| The study sessions were led by AP experts who taught NMSI-created lessons | 3.32 (0.75) | 3.53 (0.56) | -0.207 | 0.102 | -2.032 | 113.84 | **0.045** |
| The study sessions helped to increase student confidence | 3.19 (0.71) | 3.33 (0.69) | -0.134 | 0.103 | -1.298 | 144.42 | 0.196 |
| The study sessions highlighted the instructional needs of the students for me to continue addressing in class | 3.10 (0.91) | 3.28 (0.67) | -0.186 | 0.122 | -1.526 | 114.98 | 0.13 |
| I learned a great deal from watching the expert teachers during the student study sessions | 2.79 (1.05) | 3.04 (0.88) | -0.251 | 0.148 | -1.699 | 124.95 | 0.092 |
| I was able to take the strategies I saw employed during the study sessions back to my own classrooms to improve student achievement | 2.93 (0.95) | 3.143 (0.87) | -0.213 | 0.136 | -1.566 | 134.68 | 0.12 |

*Note.* 1 = *Strongly Disagree*, 2 = *Disagree Somewhat*, 3 = *Agree Somewhat*, 4 = *Strongly Agree*.

### 2. Mentoring: 2017–2018 Teacher Survey

Teachers were asked if mentoring was offered to them through the CRP and 149 teachers (78%) indicated it was. Of this group, 99 teachers indicated they had some contact with the mentor. We asked this subset of teachers the extent of their agreement with a set of statements aligned to the mentoring objectives. Table 25 presents results for the mentoring-related questions. Teachers had the highest agreement relating to the preparedness of their mentor ($M$ = 3.38, $SD$ = 0.92), followed by agreement that the mentor helped improve their content

knowledge (*M* = 2.96, *SD* = 0.98). Across all measures, teachers in the second year of the CRP rated the mentoring program more highly than teachers in the first year. Figure 6 highlights three of the metrics for which the difference between teacher ratings based on year of participation was greater. The most common forms of support from mentors were providing resources (53%) and supplying information on pacing and sequencing of lessons (46%).

Table 25

*Teacher Opinions of Mentoring*

| Mentoring-related statements | *n* | Mean | *SD* |
|---|---|---|---|
| The mentor was well-prepared | 71 | 3.38 | 0.92 |
| The mentor improved my content knowledge | 72 | 2.96 | 0.98 |
| Because of mentoring I am more effective | 71 | 2.83 | 1.05 |
| The mentor honed my skills and techniques | 71 | 2.77 | 1.01 |

*Note.* 1 = *Strongly Disagree*, 2 = *Disagree Somewhat*, 3 = *Agree Somewhat*, 4 = *Strongly Agree*.

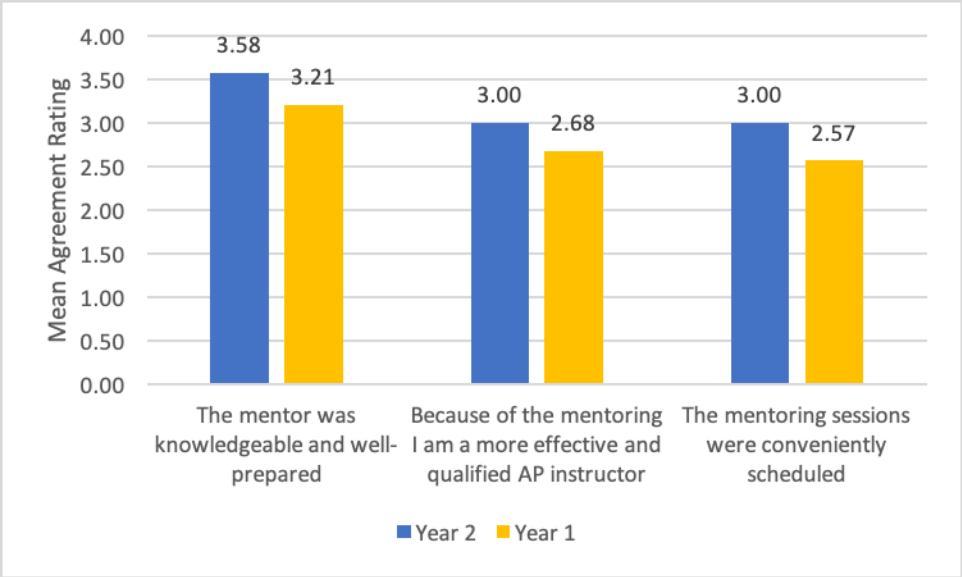



Figure 6. Teacher evaluations of mentor component by teachers' years in CRP.

### 3. Incentives: 2017–2018 Teacher Survey

Eighty-nine percent of teachers (*n* = 177) indicated they were offered incentives for teaching AP courses through the CRP. A similar number of teachers (*n* = 179) expected to receive a stipend for participating in CRP activities, such as student study sessions. Incentives were most commonly offered for students passing the AP exam (93%). More than half (52%) of teachers said the incentives were somewhat to extremely important in encouraging them to teach

AP courses (*M* = 1.89 on a 4-point scale, where 4 = *extremely important*, 3 = *important*, 2 = *somewhat important*, and 1 = *not at all important*).

### 4. Online Curricular Resources: 2017–2018 Teacher Survey

We asked teachers which additional resources were offered for use in their AP courses. Ninety-seven percent of teachers indicated they were offered access to online resources, and 69% said that they accessed the online materials on a monthly basis or more frequently. Twelve teachers indicated they did not access the online materials at all. Teachers reported using CRP materials most commonly to help familiarize students with the types of questions on the AP exam (*n* = 162). The next most commonly selected use of materials was deepening instruction in specific content areas (*n* = 143) and conducting practice exams (*n* = 137). See Figure 7.

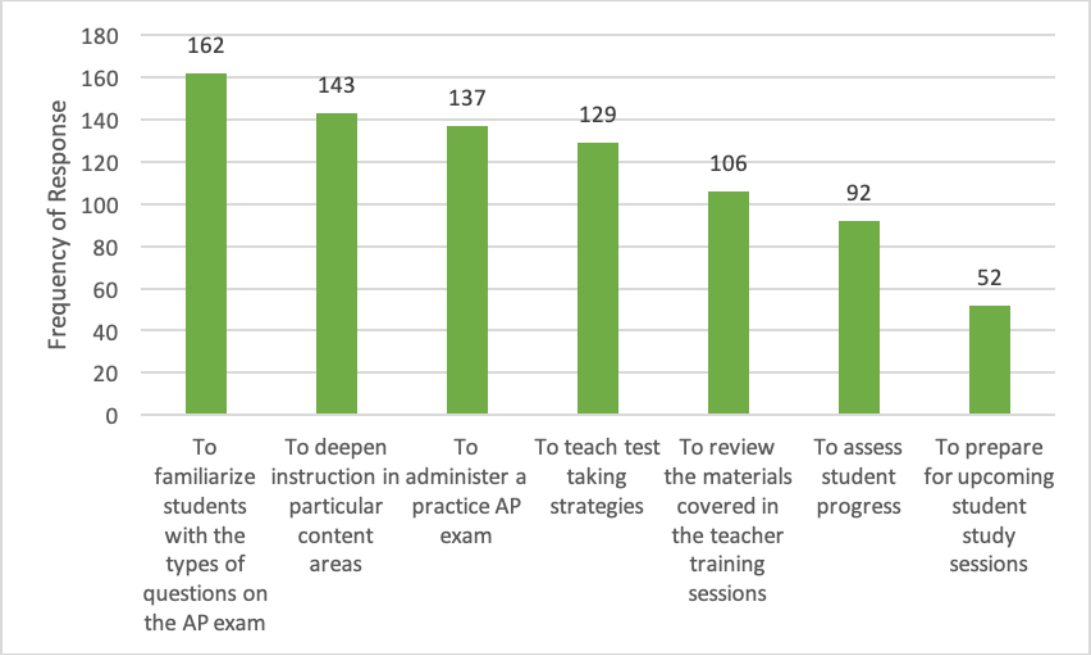



Figure 7. Teacher use of online resources.

When asked what other tools or materials they would like NMSI to provide, 56% of survey respondents selected lesson planning and 47% would like NMSI to be involved with online teacher collaboration (see Table 26).

Table 26

*Additional Tools and Materials Would Like to Have Access to From NMSI*

| Additional Tools & Materials Teachers Would Like from NMSI | Frequency | Percent |
|---|---|---|
| Lesson planning | 112 | 56% |
| Online collaboration and networking with other AP teachers | 94 | 47% |
| Online collaboration and networking with AP experts | 80 | 40% |
| Customized coaching | 47 | 23% |
| Advocacy training | 30 | 15% |
| Other | 24 | 12% |

*Note.* Multiple choice, multiple selection.

## 5. Goal Setting: 2017–2018 Teacher Survey

We asked teachers for their impression of the goals established through the CRP. Table 27 presents results from the series of goal-related questions. Nearly the same number of teachers agreed somewhat or strongly agreed that the CRP established goals for equitable access ($n$ = 158) and established measurable and attainable goals for enrollment ($n$ = 154). See Figure 8.

Table 27

*Teacher Perspectives on Goal Setting Meeting Impact and Efficacy*

| Evaluations of goal setting | *n* | *M* | *SD* |
|---|---|---|---|
| The program established goals for providing equitable access to AP coursework for all interested students | 197 | 3.05 | 0.79 |
| The program established measurable and attainable goals for class enrollment | 198 | 2.95 | 0.78 |
| The program established measurable and attainable goals for student exam performance | 198 | 2.89 | 0.85 |
| The program established goals for recruitment of high-need and traditionally underrepresented students | 198 | 2.77 | 0.86 |

*Note.* 1 = *Strongly Disagree*, 2 = *Disagree Somewhat*, 3 = *Agree Somewhat*, 4 = *Strongly Agree*.

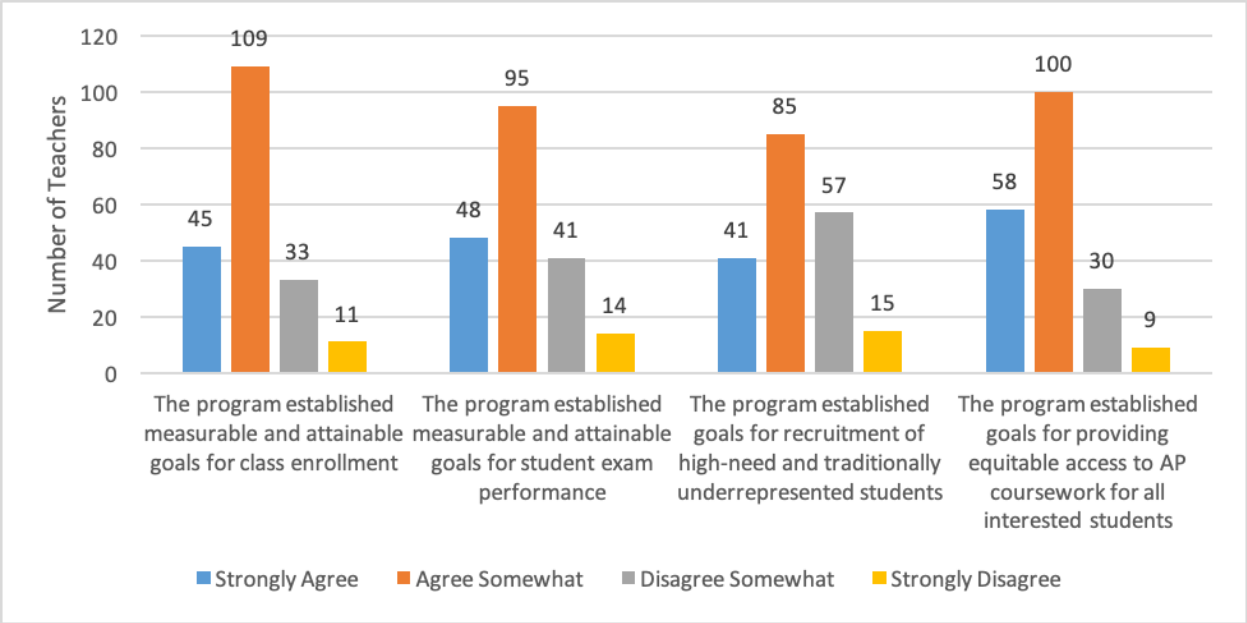



Figure 8. Teacher evaluation of goal setting.

## 6. Overall Impact of the College Readiness Program on Enrollment and Achievement: 2017–2018 Teacher Survey

Teachers were asked the extent to which they agreed with statements about the CRP's impact on student access to and success in AP math, science and English courses. Most survey participants agreed somewhat or strongly agreed with the statements (see Table 28 and Figure 9).

Table 28

*Teacher Evaluation of CRP Role in Increasing Student AP Participation*

| Role of CRP in student AP access and success | *n* | Mean | *SD* |
|---|---|---|---|
| I believe that NMSI's CRP has played an essential role in helping the school increase student success in AP math, science and English courses. | 197 | 3.30 | 0.75 |
| I believe that NMSI's CRP has played an essential role in helping the school increase student access to AP math, science and English courses. | 197 | 3.29 | 0.74 |

*Note.* 1 = *Strongly Disagree*, 2 = *Disagree Somewhat*, 3 = *Agree Somewhat*, 4 = *Strongly Agree*.

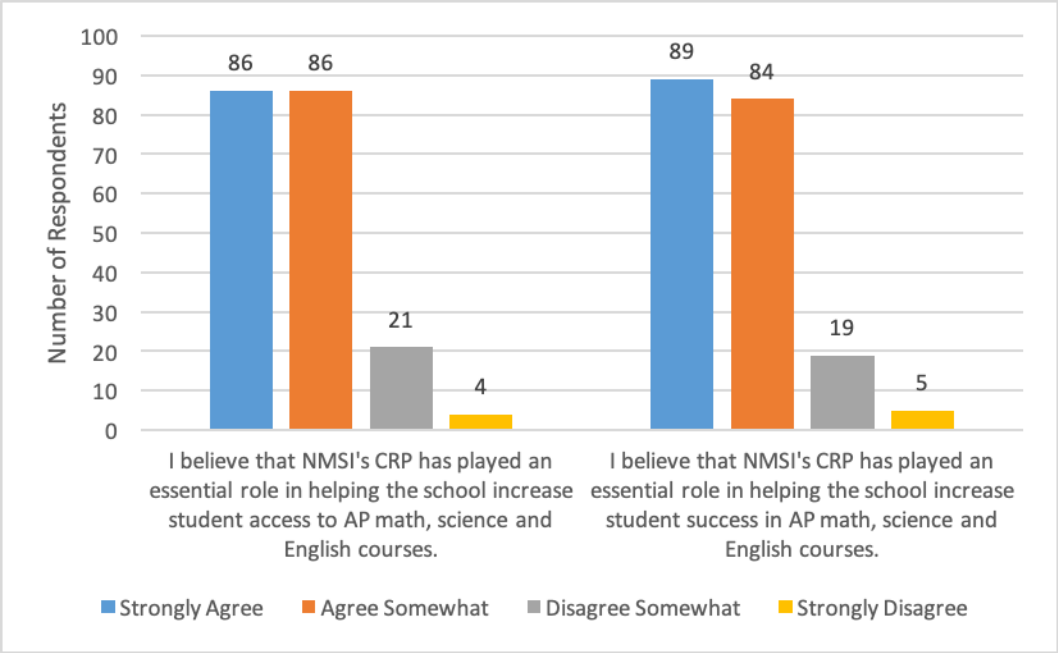



Figure 9. Teacher evaluation of CRP role in increasing student AP participation.

However, when we compared respondents from Treatment Schools to those from Delayed Treatment Schools, we discover that—even though both groups believe that NMSI contributed to increasing student access to AP courses—Treatment School teachers felt significantly more strongly about the positive impact (see Table 29). This may be related to Treatment School teachers also feeling more strongly that their schools encouraged all students to enroll in AP courses (see Table 30).

Table 29

*CRP Role in Increasing Student Access to AP*

| **Increase student access to AP** | **Delayed Treatment *M (SD)*** | **Treatment *M (SD)*** | **DT - T diff** | **S.E.** | ***t*** | **Df** | ***p*** |
|---|---|---|---|---|---|---|---|
| I believe that NMSI's CRP has played an essential role in helping the school increase student access to AP math, science and English courses. | 3.11 (0.75) | 3.39 (0.72) | -0.276 | 0.109 | -2.529 | 140.012 | 0.013 |

*Note.* 1 = *Strongly Disagree*, 2 = *Disagree Somewhat*, 3 = *Agree Somewhat*, 4 = *Strongly Agree*.

Table 30

*School Encourages All Students to Enroll in AP*

| Encouraging enrollment | Delayed Treatment *M (SD)* | Treatment *M (SD)* | DT - T diff | S.E. | *t* | Df | *p* |
|---|---|---|---|---|---|---|---|
| My school encourages all students to enroll in AP courses. | 2.86 (0.84) | 3.14 (0.88) | -0.281 | 0.126 | -2.219 | 152.661 | 0.028 |

*Note.* 1 = *Strongly Disagree*, 2 = *Disagree Somewhat*, 3 = *Agree Somewhat*, 4 = *Strongly Agree*.

## 7. Student Enrollment in AP Courses: 2017–2018 Teacher Survey

We asked teachers if they felt that the CRP was an effective way to increase student enrollment in AP courses, and 83% of teachers ($n$ = 165) reported that it was. Several teacher comments focused on the classroom challenges of open enrollment, such as "all you are doing is watering down AP courses." We also asked teachers if opening up enrollment to all students had a positive impact on the AP program, and 81% of teachers said that it had ($n$ = 161).

Thirty percent of teachers said that the CRP contributed to a major improvement in recruitment of high-need and traditionally underrepresented students into AP course, and 51% of teachers felt it contributed to at least a slight improvement in this area ($n$ = 198). In the 2018 survey, we saw a significant increase (among the 83 teachers who completed surveys both years) in the percentage who felt that open enrollment had a positive imact on the AP program at their school (see Table 31).

Table 31

*Whether Opening Up Enrollment to All Students Had a Positive Impact on the AP Program*

| Positive impact on AP program of open enrollment | 2017 *M (SD)* | 2018 *M (SD)* | 2017 - 2018 diff | S.E. | t | Df | p |
|---|---|---|---|---|---|---|---|
| Yes = 1 | 0.65 (0.48) | 0.82 (0.39) | -0.17 | 0.064 | -2.188 | 70 | 0.032 |

To further explore factors related to enrollment, we asked teachers in Delayed Treatment Schools if their school offered AP courses in the 2016–2017 school year. Teachers from each of the 21 Delayed Treatment Schools verified that their schools offered AP courses prior to implementing the CRP. Teachers were then asked to select the most important reasons, from a list of ten reasons (plus other), why qualified students may not have enrolled in AP courses in the past. The most common reason given was that students have too many scheduling conflicts, which was selected by 65% of the teachers responding to the item ($n$ =108). Other options

frequently chosen were that the courses have a reputation as being difficult (57%) and that students prefer to enroll in dual credit classes with local colleges (41%). Sixteen teachers provided additional comments for this question, citing student apprehension at the increased work load, student concern about jeopardizing their GPA, and students opting for International Baccalaureate courses instead of AP courses as common reasons students have previously not enrolled in AP courses.

We also asked survey participants from Delayed Treatment Schools in 2018 if opening up enrollment was a benefit to the AP program at their school, and compared their responses to those of Treatment School teachers in the 2016–2017 school year. Table 32 reveals that a significantly greater percentage of Delayed Treatment teachers (80%) thought that open enrollment had a positive impact on the AP program (compared to 60% in the Treatment Schools). This could be related to the number of schools in the group with an extant AP program. If the Treatment Schools had fewer sites with existing AP programs, then there would not be anything to improve.

Table 32

*Whether Open Enrollment Was a Benefit the First Year of the Program*

| **Positive impact on AP program of open enrollment** | **Treatment 2017** ***M (SD)*** | **Delayed Treatment 2018** ***M (SD)*** | **2017 - 2018 diff** | **S.E.** | **t** | **df** | **p** |
|---|---|---|---|---|---|---|---|
| Yes=1 | 0.6 (0.492) | 0.803 (0.401) | -0.203 | 0.066 | -3.069 | 170.346 | 0.002 |

## 8. Student Performance in AP Courses

We compared the first year experiences of teachers during the first year the program was implemented in their school (i.e., 2016–2017 school year for Treatment Schools, $n = 129$, and 2017–2018 school year for Delayed Treatment Schools, $n = 72$). We would expect there to be, if the groups are roughly equivalent, few differences between the groups. Table 33 shows however, that, on average, teachers in Delayed Treatment Schools believe their students were better prepared to take AP courses and AP exams. Again, this may be related to the fact that the Delayed Treatment Schools had more extant AP courses prior to participation in the CRP than did the Treatment Schools.

Table 33

*Student Preparedness the First Year of Program Implementation*

| Statements about student preparedness | Treatment 2017 *M (SD)* | Delayed Treatment 2018 *M (SD)* | 2017 - 2018 diff | S.E. | t | df | *p* |
|---|---|---|---|---|---|---|---|
| Students in AP courses in my school believe that they are well-prepared for the exam. | 2.659 (0.643) | 2.903 (0.653) | -0.244 | 0.096 | -2.551 | 145.065 | 0.012 |
| Students in AP courses in my school are well-prepared for the exam. | 2.625 (0.71) | 2.875 (0.691) | -0.25 | 0.103 | -2.432 | 150.691 | 0.016 |
| The students in my school are well-prepared to take AP courses. | 2.434 (0.799) | 2.75 (0.835) | -0.316 | 0.121 | -2.612 | 141.584 | 0.01 |

*Note.* 1 = *Strongly Disagree*, 2 = *Disagree Somewhat*, 3 = *Agree Somewhat*, 4 = *Strongly Agree*.

### 9. Effective Components of the CRP

From a list of the many elements of the CRP, 54% of participants chose teacher training as the most effective component of the CRP (*n* = 108). When asked to select the *second* most effective component of the program, the top two choices were teacher training and the provision of classroom equipment and materials, both selected by 23% of 200 teachers. Mentoring was the most commonly selected component seen as the *least* effective, chosen by 37% of 194 teachers (see Figure 10).

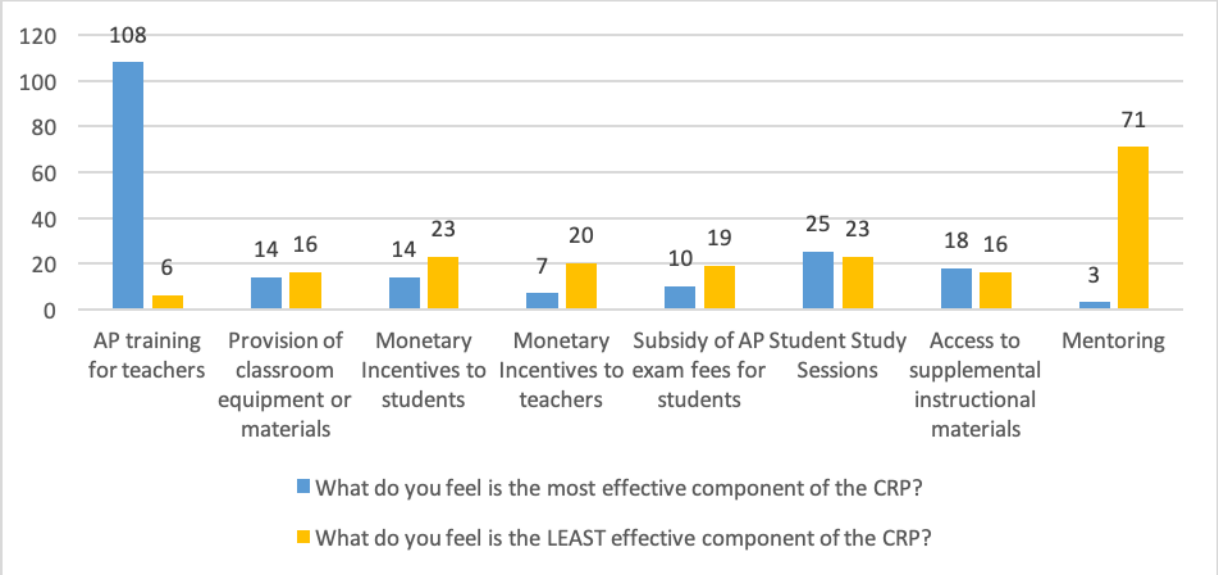



Figure 10. Comparison of teachers' most/least effective program elements.

The 2018 survey participants from the Treatment Schools, when compared to respondents from Delayed Treatment Schools, believed more strongly that the CRP contributed to improvements in student experience with STEM AP courses, teacher content knowledge, teacher instructional skills and strategies, the school culture of continuous improvement, and school leadership valuing STEM learning (see Table 34).

Table 34

*Improvements Attributable to NMSI*

| **Degree to which NMSI improved the following** | **Delayed Treatment *M (SD)*** | **Treatment *M (SD)*** | **DT - T diff** | **S.E.** | **t** | **Df** | ***p*** |
|---|---|---|---|---|---|---|---|
| Students' content knowledge | 2.36 (0.66) | 2.42 (0.58) | -0.056 | 0.093 | -0.604 | 133.681 | 0.547 |
| Students' experience with STEM AP courses | 2.18 (0.66) | 2.38 (0.59) | -0.193 | 0.095 | -2.038 | 132.53 | **0.044** |
| Recruitment of high-need and traditionally underrepresented students into AP courses | 1.97 (0.69) | 2.18 (0.69) | -0.203 | 0.102 | -1.979 | 148.303 | 0.05 |
| Teachers' content knowledge | 2.32 (0.69) | 2.59 (0.55) | -0.275 | 0.095 | -2.896 | 122.887 | **0.004** |
| Teachers' instructional skills, techniques and strategies | 2.39 (0.69) | 2.62 (0.55) | -0.223 | 0.095 | -2.351 | 120.141 | **0.02** |
| School culture of continuous improvement | 2.01 (0.64) | 2.21 (0.66) | -0.191 | 0.095 | -2.002 | 151.377 | **0.047** |
| School leadership valuing STEM learning | 2.03 (0.67) | 2.24 (0.64) | -0.211 | 0.099 | -2.122 | 132.663 | **0.036** |

*Note.* 1 = *No Improvement*, 2 = *Slight Improvement*, 3 = *Major Improvement*.

### a. 2017–2018 Partner School Director Survey Results

Next we present findings from the Partner School Director survey. CRP administrators were asked their opinions on the effectiveness of the CRP. In general, Partner School Directors in Treatment Schools felt more strongly that the CRP met their expectations which was also reflected in more positive opinions about the efficacy of the program (see Figure 11 and Table 35).

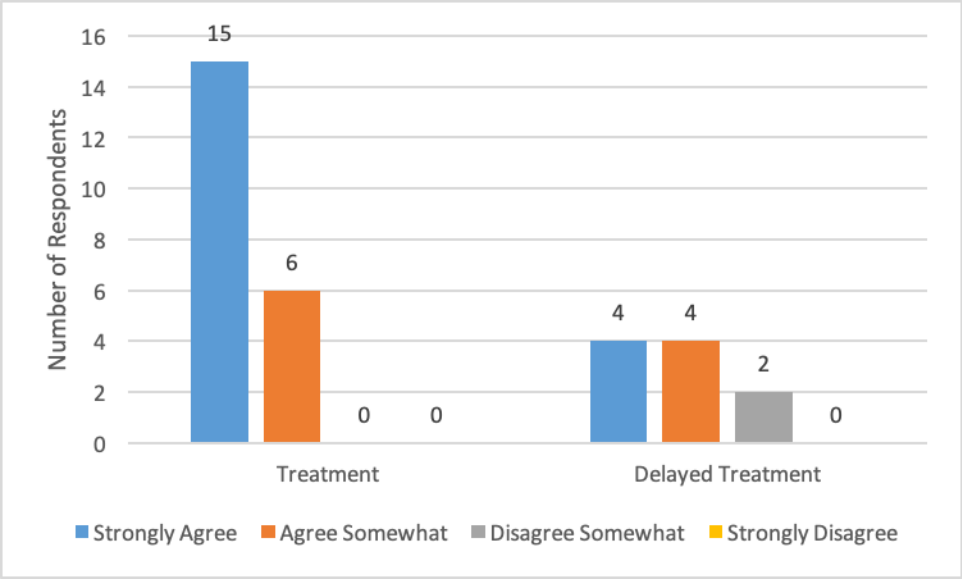



Figure 11. Whether CRP matched PSD expectations in 2017–2018 school year.

Table 35

*PSD Evaluation of CRP Influence on Culture by School Group*

| | Treatment | | | Delayed Treatment | | |
|---|---|---|---|---|---|---|
| **NMSI influence on culture** | ***N*** | ***M*** | ***SD*** | ***N*** | ***M*** | ***SD*** |
| NMSI provided a platform for networking and collaboration. | 21 | 3.76 | 0.43 | 10 | 3.50 | 0.67 |
| I have the support, resources and training to successfully increase access and success in math, science and English AP courses. | 21 | 3.62 | 0.49 | 10 | 3.50 | 0.50 |
| I believe that NMSI's CRP has played an essential role in helping the school increase student success in AP math, science and English courses. | 21 | 3.62 | 0.49 | 10 | 3.20 | 0.60 |
| I believe that NMSI's CRP has played an essential role in helping the school increase student access to AP math, science and English courses. | 21 | 3.52 | 0.59 | 10 | 3.20 | 0.60 |
| I believe that with proper support any student in this school can take an AP course and be successful. | 21 | 3.48 | 0.73 | 10 | 2.90 | 0.83 |

*Note.* 1 = *Strongly Disagree*, 2 = *Disagree Somewhat*, 3 = *Agree Somewhat*, 4 = *Strongly Agree*.

A follow-up question focused specifically on open enrollment. Ninety-seven percent of Partner School Directors (32/33) said open enrollment had a positive impact on the AP program at their school and that the CRP was an effective way to increase student enrollment in AP courses. Forty-two percent of respondents agreed that the CRP contributed to a major improvement in recruitment of high-need and traditionally underrepresented students into AP courses, and 58% felt it contributed to a slight improvement in this area.

An additional set of questions asked administrators to indicate the extent to which the CRP contributed to improvements in certain areas. The highest level of perceived improvement was in teachers' instructional skills, techniques, and strategies (*M* = 2.73, *SD* = 0.45), followed by teachers' content knowledge (*M* = 2.61, *SD* = 0.49). Seventy-three percent of respondents (*N* = 33) thought that the CRP contributed to a major improvement in teachers' instructional skills, and 61% indicated major improvement in teachers' content knowledge. School leadership valuing STEM learning was rated as less impacted by CRP, with 52% of respondents indicating a slight improvement (*M* = 2.30, *SD* = 0.63). In all cases, however, the average impact was at least slight improvement (see Table 36).

Table 36

*Administrator Perceptions of CRP-Related Improvements*

| Areas of Improvement | *n* | Mean | *SD* |
|---|---|---|---|
| Teachers' instructional skills, techniques and strategies | 33 | 2.73 | 0.45 |
| Teachers' content knowledge | 33 | 2.61 | 0.49 |
| Students' content knowledge | 33 | 2.52 | 0.50 |
| Recruitment of high-need and traditionally underrepresented students into AP courses | 33 | 2.42 | 0.49 |
| Students experience with STEM AP courses | 33 | 2.36 | 0.64 |
| School culture of continuous improvement | 33 | 2.36 | 0.54 |
| School leadership valuing STEM learning | 33 | 2.30 | 0.63 |

*Note.* 1 = *No Improvement*, 2 = *Slight Improvement*, 3 = *Major Improvement*.

### b. 2017–2018 Teacher and PSD Interviews

We interviewed 85 teachers from 24 schools in seven states as well as 14 administrators from 14 schools. This group of teachers represents a convenience sample of schools and teachers based on availability and scheduling constraints. Most teachers were interviewed at their schools, and some were interviewed during summer training sessions. Interview questions were drawn from and aligned to key themes of the online survey, but were open ended to encourage less restricted conversation.

#### 1. 2017–2018 Interviews: Overall Efficacy of the CRP

We first asked teachers their opinions on the most important or effective component of the CRP. Thirty-eight percent of the teacher interviewee sample indicated that the training was the most effective component, which mirrored the findings from the teacher surveys. As in 2016–2017, the second most frequently stated component was resources (22%) followed by student

study sessions (16%). Teachers often included additional effective elements in their responses. Considering all components cited by interviewees as effective (as opposed to just the one mentioned as the *most* effective), 73% of teachers indicated that training was an effective component. Administrators also most frequently listed teacher training as the most effective component of the program. However, second year PSDs were just as likely to list changes in their school's culture as the most important component of the CRP.

Sixty-four percent of teachers and 77% of administrators felt that school culture had changed since the implementation of the CRP. In 2016–2017, almost half of the interviewees who felt the culture had not changed (7 of 15) said that it was most likely too soon to tell what impact the program would have on the school culture. In 2017–2018, second year teachers in the program and teachers in Treatment Schools were more likely to claim that school culture had changed (see Table 37).

Table 37

*Teacher Perspective on School Culture Change*

| Statement | 1st Year CRP Teacher | 2nd Year CRP Teacher | Teacher in Delayed Treatment School | Teacher in Treatment School |
|---|---|---|---|---|
| % Saying Culture Changed | 54% | 81% | 54% | 73% |

### 2. 2017–2018 Interviews: Specific CRP Components

The majority of teachers (85%) were satisfied with the level of training and support received throughout the academic year as part of the CRP, and 97% said they felt "adequately prepared" to teach their AP course. Ninety-four percent of teachers were offered a mentor, but only 27% of those teachers took advantage of the opportunity to meet with the mentor. Of those who took advantage of the mentor opportunity, 95% said it was beneficial.

We asked teachers if they would change anything about the CRP to help improve AP education at their school. Twenty-two percent of teachers had suggestions for improving the training sessions, and 19% had suggestions about the student study sessions.

### c. 2017–2018 Student Survey Results

Below we summarize data from the 1,930 students who responded to the CRP student survey (see Table 38 for the distribution of respondents). Surveys were sent to teachers who, through an item on the teacher survey, expressed an interest in administering them. Students completed paper-and-pencil copies of the survey which were then mailed back.

Table 38

*Distribution of Student Survey Respondents by State*

| State | *n* | % |
|---|---|---|
| CA | 128 | 7% |
| GA | 217 | 11% |
| IL | 253 | 13% |
| LA | 85 | 4% |
| MI | 131 | 7% |
| MO | 120 | 6% |
| ND | 225 | 12% |
| OH | 174 | 9% |
| PA | 357 | 18% |
| TX | 240 | 12% |

### 1. 2017–2018 Student Survey: AP Courses

The average number of computer science, math, science, and English AP courses taken by all students in the sample was two (*SD* = 0.97, *Mdn* = 2.00), with a range from 1-6. Figure 12 shows the difference in course loads between students in Treatment and Delayed Treatment Schools.

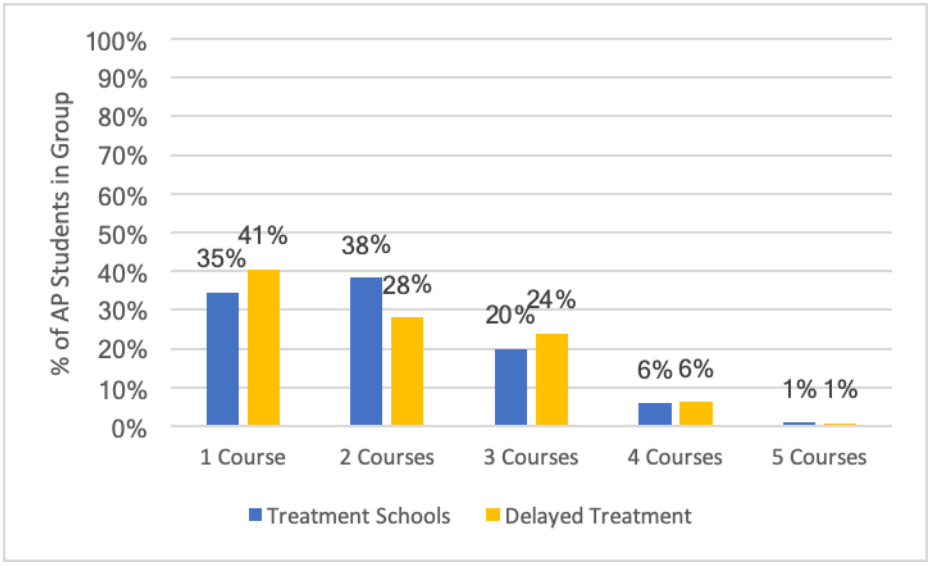



Figure 12. Student self-report number of STEM/ELA AP courses.

Figure 13 shows the percentage of students who reported taking specific AP courses during 2017–2018. The largest number of respondents were taking AP English Language (*N* = 941; 49%), followed by calculus (*N* = 532; 28%).

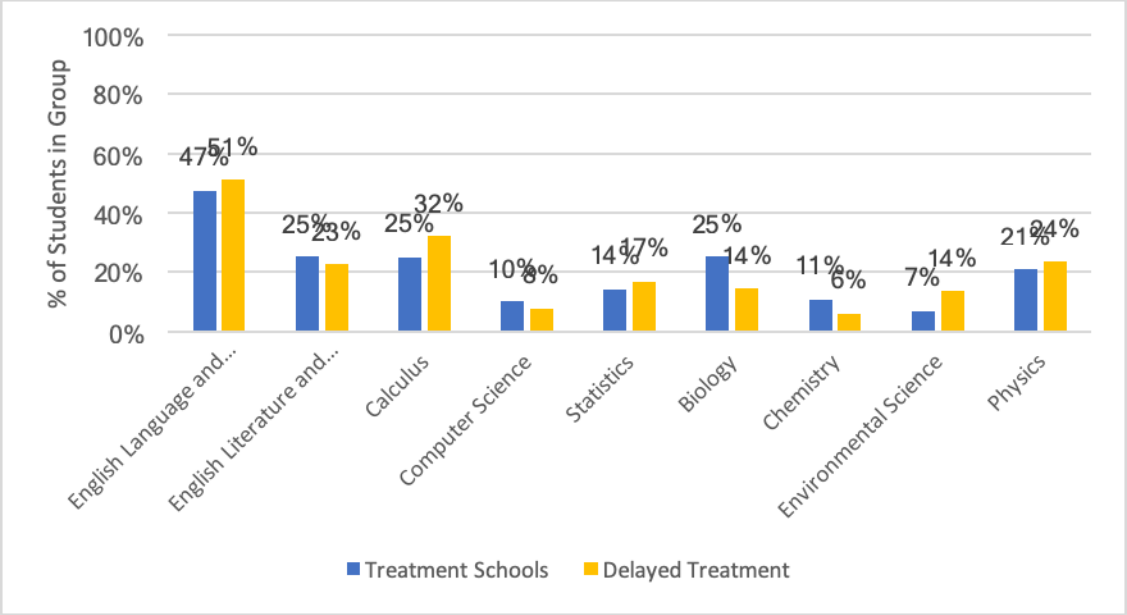



Figure 13. Percentage of participants taking each course.

### 2. 2017–2018 Student Survey: Knowledge of AP and the CRP

We asked students how they learned about the AP program in their school as well as how they learned about the CRP specifically. Most students learned about AP and the CRP from their teachers or school counselors, and more than a third of students learned about the AP courses from other students (see Table 39). In terms of learning about AP courses, responses in the "other" category included online research and the school's course directory. Some students indicated that they were just automatically enrolled in the class or that the course was required. When asked how they learned about the CRP, 29% of the students who responded to the item said that they had not heard of the CRP.

Table 39

*Sources of Student Knowledge of the AP and College Readiness Program*

| Source | Learn about AP (*n* = 1,925) | Learn about CRP (*n* = 1,913) |
|---|---|---|
| My AP Teachers | 1,169 (61%) | 1,116 (58%) |
| Other teachers at school | 865 (45%) | 272 (14%) |
| School counselor | 791 (41%) | 345 (18%) |
| Other students | 824 (43%) | 195 (10%) |
| School signs, emails, fliers | 271 (14%) | 125 (7%) |
| Family members | 236 (12%) | |
| I had not heard about CRP | | 546 (29%) |
| Other | 59 (3%) | 13 (1%) |

### 3. 2017–2018 Student Survey: Future Educational Plans

Ninety-one percent of student respondents (*n* = 1,749) indicated that they plan to attend some sort of postsecondary institution. Of these the highest level of education students planned to complete was two-year community/junior college (2%), four-year college/university (43%), and graduate school (46%).

### 4. 2017–2018 Student Survey: Student Study Sessions

Survey participants were next asked a set of questions about the student study sessions. Students were required to attend three student study sessions for each AP course in which they were enrolled, although we learned through some of the teacher interviews that many students had difficulty attending the weekend sessions because of conflicting demands and scheduling issues. As Figure 14 shows, the majority of survey participants attended at least one student study session.

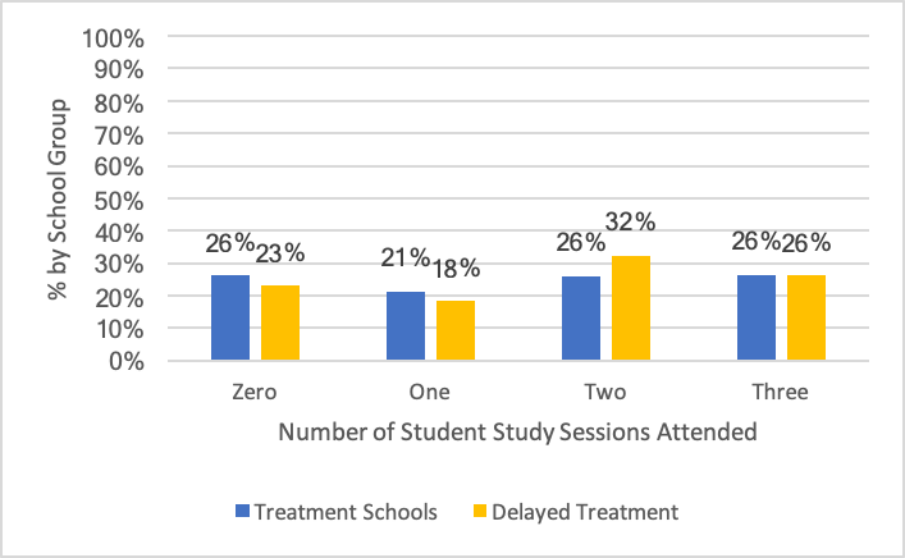



Figure 14. Number of student study sessions attended.

Students learned about the student study sessions most frequently from their AP teachers (98% of the time) as well as other teachers and school counselors. Forty-one percent of students ($n$ = 778) said that their school provided transportation to the student study sessions, and 85% ($n$ = 1,621) said that some rewards were offered for attending the sessions (the item suggested extra credit and lunch as examples of rewards). When asked if there was an opportunity to make up study sessions if they could not attend, only 52% of students answered affirmatively. Figure 15 illustrates the difference between Treatment School and Delayed Treatment School execution of student study sessions. Of the 1,495 students who responded, the average rating of usefulness of the student study sessions was 3.04 ($SD$ = 0.77), or somewhat useful. As Table 40 shows, the higher the percentage of study sessions attended, the more useful students found them.

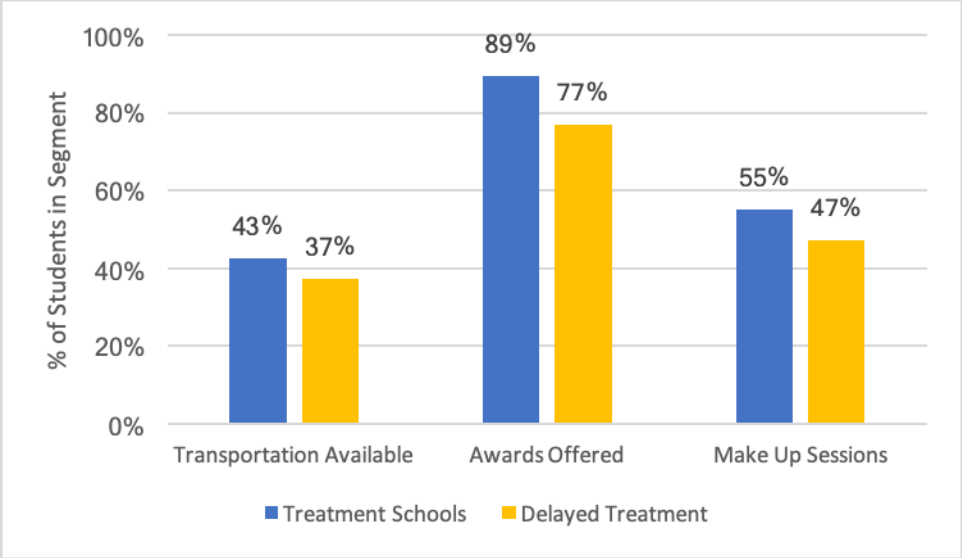



Figure 15. School efforts to support student study sessions.

Table 40

*Usefulness of Student Study Sessions*

| Number of sessions attended | *n* | Average rating |
|---|---|---|
| None | 52 | 2.69 |
| Some, but not all | 827 | 3.00 |
| All | 596 | 3.15 |

*Note.* 1 = *not at all useful*, 2 = *slightly useful*, 3 = *somewhat useful*, and 4 = *extremely useful.*

## 5. 2017–2018 Student Survey: Student Opinions about Study Sessions

We asked students how much they agreed with statements related to the study sessions, for the AP course in which they were completing the survey. Table 41 presents the average agreement of student respondents: on average, students had the lowest level of agreement that the in-person study sessions were conveniently scheduled (on a 4-point scale). The highest level of agreement was with the study sessions improving students' content knowledge. See also Figure 16, Figure 17, and Table 42).

Table 41

*Student Responses to Study Session Statements*

| Student evaluations of student study sessions | *n* | *M* | *SD* |
|---|---|---|---|
| Improved my content understanding | 1,480 | 3.12 | 0.70 |
| The study sessions increased my confidence | 1,481 | 3.01 | 0.72 |
| Improved my test-taking strategies | 1,476 | 2.96 | 0.82 |
| Increased my confidence in my ability to take the AP exams | 1,476 | 2.95 | 0.78 |
| Conveniently scheduled | 1,468 | 2.86 | 0.87 |

*Note.* 1 = *Strongly Disagree*, 2 = *Disagree Somewhat*, 3 = *Agree Somewhat*, 4 = *Strongly Agree*.

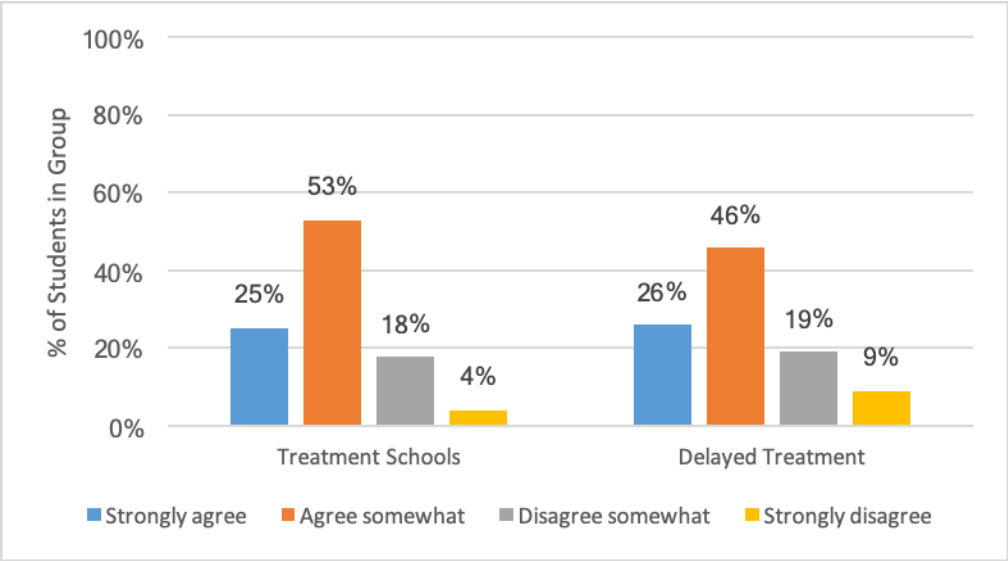



Figure 16. Student study sessions—improvement of test-taking strategies.

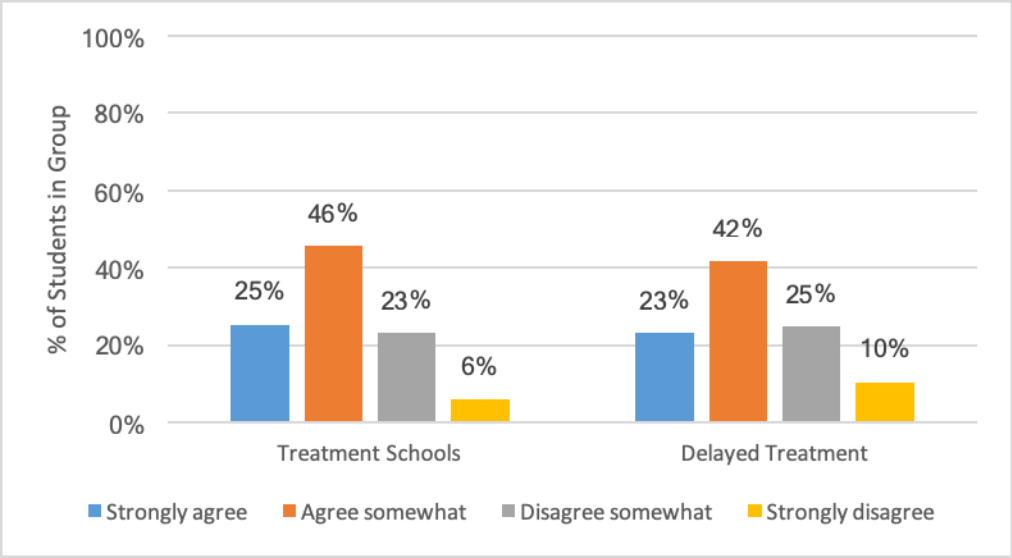



Figure 17. Student study sessions—convenience of scheduling.

Table 42

*Student Study Session Evaluations by Number of Sessions Attended*

| | All | | Some | | None | |
|---|---|---|---|---|---|---|
| **Student evaluation of student study sessions** | ***n*** | ***M*** | ***n*** | ***M*** | ***n*** | ***M*** |
| The Study Sessions improved my understanding of the course content | 590 | 3.22 | 821 | 3.07 | | |
| The Study Sessions increased my confidence in my ability to successfully complete the AP course | 590 | 3.11 | 821 | 2.97 | | |
| The Study Sessions improved my test taking strategies | 588 | 3.06 | 819 | 2.91 | | |
| The Study Sessions increased my confidence in my ability to get a score of 3 or more on the AP exam | 588 | 3.05 | 819 | 2.89 | | |
| The in-person Study Sessions were conveniently scheduled | 584 | 2.96 | 813 | 2.79 | 49 | 2.84 |

*Note.* 1 = *Strongly Disagree*, 2 = *Disagree Somewhat*, 3 = *Agree Somewhat*, 4 = *Strongly Agree*.

## 6. 2017–2018 Student Survey: AP Preparation

Students were asked to indicate their level of agreement with a set of statements related to their preparation for the AP exams (see Table 43). Lowest levels of agreement were found for statements relating to outside classroom support for increasing content understanding ($M$ = 2.86, $SD$ = 0.81) and improving test-taking strategies ($M$ = 2.84, $SD$ = 0.83). Students' highest level of agreement was related to their AP teachers' content understanding with an average agreement of 3.78 ($SD$ = 0.50). See Table 43 and Figure 18.

Table 43

*Student Opinions on Factors Relating to AP Preparation*

| Opinions about AP preparation | *n* | *M* | *SD* |
|---|---|---|---|
| My AP teachers understand the content they are teaching. | 1,901 | 3.78 | 0.50 |
| I am confident in my ability to successfully complete AP courses. | 1,895 | 3.33 | 0.70 |
| I am confident in my ability to learn new STEM content. | 1,895 | 3.26 | 0.70 |
| I am confident in my ability to successfully take AP exams. | 1,893 | 3.09 | 0.79 |
| I was nervous about how hard the AP courses would be when I signed up for them. | 1,897 | 3.05 | 0.95 |
| The support the school provides outside of the classroom improved my study skills. | 1,902 | 2.89 | 0.81 |
| The support the school provides outside of the classroom increased my understanding of the course content. | 1,891 | 2.86 | 0.81 |
| The support the school provides outside of the classroom improved my test taking strategies. | 1,898 | 2.84 | 0.83 |

*Note.* 1 = *Strongly Disagree*, 2 = *Disagree Somewhat*, 3 = *Agree Somewhat*, 4 = *Strongly Agree*.

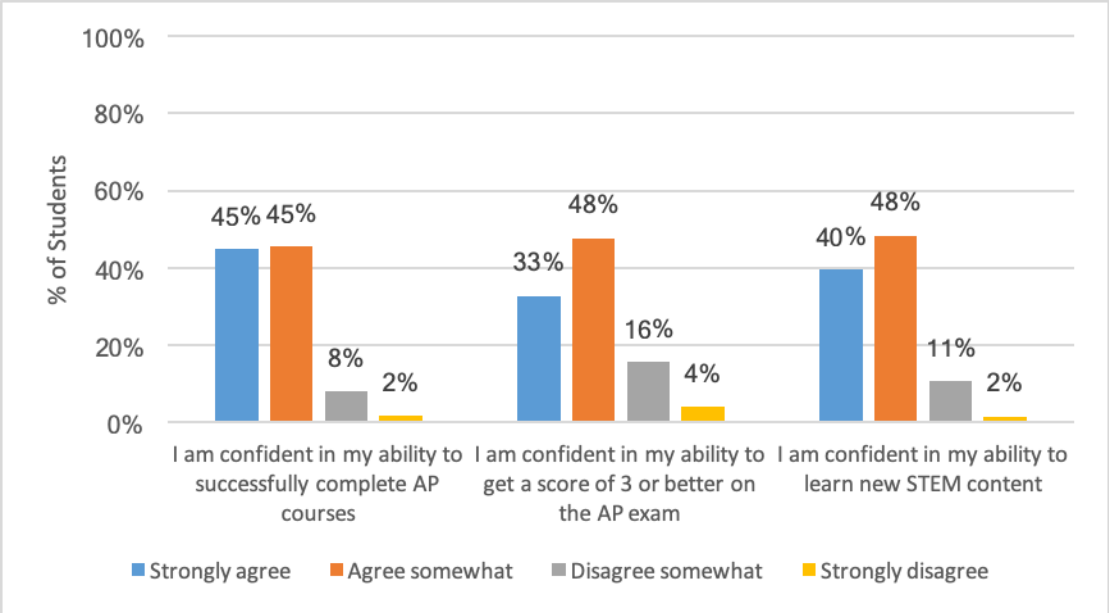



Figure 18. Self-assessment of students' confidence.

For nearly every item addressing student confidence or the benefits of the supports students received outside of the classroom, the more student study sessions a student attended the more likely they were to agree with the statements. The lone exception was the statement, "I am confident in my ability to learn new STEM content." Despite that distinction, the item remains

one of the top two rated items in all categories (i.e., no sessions attended, some attended, and all attended). See Table 44.

Table 44

*AP Preparedness by Number of Student Study Sessions Attended*

| | All | | Some | | None | |
|---|---|---|---|---|---|---|
| **Evaluation of AP preparedness** | ***n*** | ***M*** | ***n*** | ***M*** | ***n*** | ***M*** |
| I am confident in my ability to successfully complete AP courses | 585 | 3.38 | 823 | 3.32 | 414 | 3.31 |
| I am confident in my ability to learn new STEM content | 589 | 3.24 | 819 | 3.23 | 413 | 3.31 |
| I am confident in my ability to get a score of 3 or better on the AP exam | 588 | 3.11 | 819 | 3.09 | 413 | 3.08 |
| The support the school provides outside of the classroom increased my understanding of the course content | 585 | 3.01 | 819 | 2.82 | 413 | 2.72 |
| The support the school provides outside of the classroom improved my study skills | 589 | 2.99 | 825 | 2.84 | 415 | 2.84 |
| The support the school provides outside of the classroom improved my test taking strategies | 588 | 2.97 | 823 | 2.81 | 413 | 2.70 |

*Note.* 1 = *Strongly Disagree*, 2 = *Disagree Somewhat*, 3 = *Agree Somewhat*, 4 = *Strongly Agree*.

### 7. 2017–2018 Student Survey: AP Exams

Only 98 students indicated they were not planning on taking the AP exam for the course in which they were completing the survey. We asked these students why they were deciding not to take the exam and a summary of responses are shown in Figure 19. The most common response was that the student did not feel ready to take the exam, with the second-most common response having a similar theme (course load was too heavy and so they did not feel they could prepare).

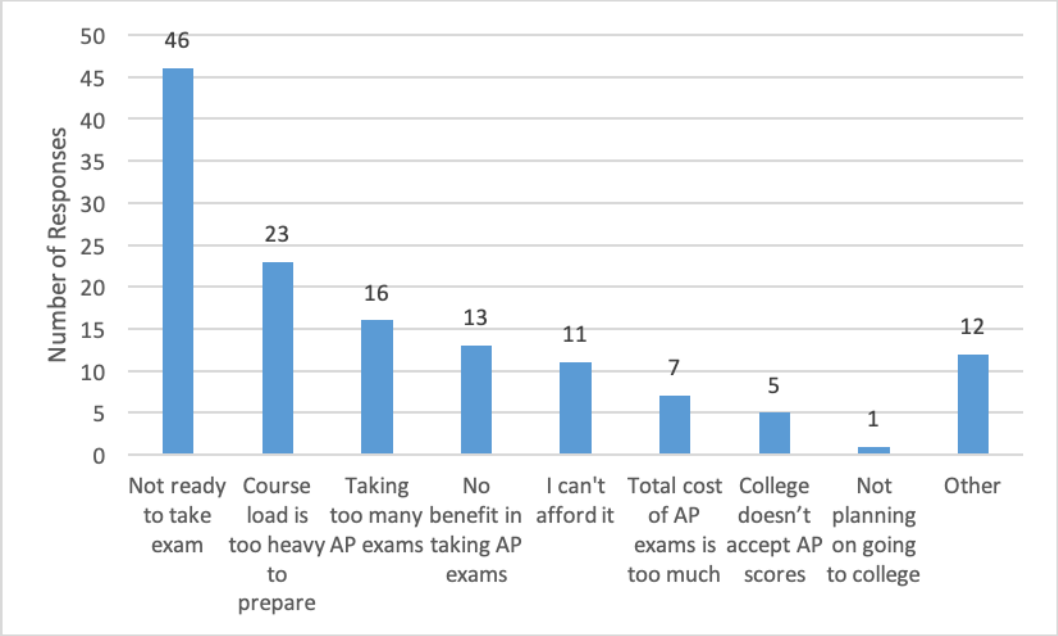



Figure 19. Reasons given by students for not taking the AP exam.

### 8. 2017–2018 Student Survey: Student Incentives and Rewards

Students were asked which rewards or incentives were offered to them to participate in AP courses. Sixty-nine percent of students were offered cash awards and 47% indicated they had their AP exam fee waived. Weighted grades (63%) and dual credit in a local college (38%) were also cited as incentives. To earn the cash rewards, 82% of the students indicated they had to pass the AP exam (with a score of 3 or higher), 26% had to complete and pass an AP course, and 24% said they had to take the AP exam. On a four-point scale students rated the importance of the incentives an average of 2.66 (*SD* = 1.02), indicating that the cash incentives were slightly important.

## D. 2018–2019 Fidelity Matrix Data Collection

### a. 2018–2019 Teacher and Administrator Surveys

As in previous years of the study, most online survey questions were multiple choice, multiple select (respondents could choose more than one answer), or questions with a Likert scale (most often four-point). Most items were consistent with the 2018 survey, and many items remained unchanged from the 2017 surveys.

### b. 2018–2019 Student Survey

Student surveys were distributed in the same way as in the previous year.

### c. 2018–2019 Teacher and Administrator Interviews

Also as in previous years, we conducted interviews with a subset of CRP teachers and administrators to supplement information from the surveys.

## E. 2018–2019 Fidelity Matrix Results

In the third year of the study, we once again utilized a fidelity matrix to evaluate the implementation of the CRP in 48 schools in 10 states. For the 2018–2019 school year matrix, 19 measures examined the level of participation of school personnel and students as well as NMSI's fulfillment of various program-related administrative responsibilities.

Data to complete the matrix were gathered from administrative records as well as survey and interview responses where necessary. In the spring of 2019, 211 teachers completed online surveys including at least one survey from each of the 48 schools being evaluated (see Table 45). Partner School Directors (PSDs), Site Coordinators (SCs), and students also completed surveys. Additionally, researchers interviewed PSDs and/or teachers from 31 out of 48 schools.

Table 45

*2018–2019 Survey and Interview Participation*

| Instrument | Number completed | Schools represented |
|---|---|---|
| Teacher Survey | 211 | 48 |
| Partner School Director Survey | 28 | 28 |
| Site Coordinator Survey | 36 | 34 |
| Student Survey | 2,710 | 38 |
| Teacher Interview | 95 | 31 |
| Partner School Director Interview | 11 | 11 |

### a. School Indicators: 2018–2019 Fidelity Matrix

All seven "school" matrix components evaluate NMSI's fulfillment of specific administrative responsibilities, including Program Manager support. NMSI outlines four key responsibilities for their school liaisons known as Program Managers:

- delivery of student study session reminders and materials
- assistance with participant registration for teacher training and student study sessions
- assignment of mentor teachers
- guidance about locating and incorporating NMSI resources into the curriculum

Survey respondents were asked directly if the Program Manager performed each of the four functions. Both the teacher and PSD surveys offered participants a simple yes/no choice. Because the responsibilities and program involvement of the Site Coordinator varied from school to school, the Site Coordinator survey item included a "don't know" option.

The first two components of Program Manager support are school-wide actions that could be observed by any survey participant, so responses to all three surveys were considered. For the purposes of the fidelity matrix calculation, a school was deemed in compliance separately for each of the first two components of Program Manager support if a participant in any of the three CRP roles answered affirmatively.

The latter two components could be true or false on a teacher-by-teacher basis, and only the teacher responses were considered when determining fidelity. As with the other Program Manager elements, fidelity was achieved by at least one survey participant affirming the support. While NMSI was viewed as compliant at many schools through these metrics (see Table 46), it provides a more rounded view to note that 31% of teachers responding to the item said they did not get support matching to a mentor and 39% said they did not get support with locating NMSI resources. To be clear, a survey with no response was not considered a negative response.

Forty-one schools indicated that NMSI Program Managers fulfilled all four functions, and the remaining seven schools indicated support in three categories.

Table 46

*2018–2019 Program Manager Support*

| Area of Program Manager support | Number of schools |
|---|---|
| Delivery of student study session reminders and materials | 48 |
| Assistance in teacher sign-ups for training and student study sessions | 48 |
| Assignment of mentor teachers | 45 |
| Guidance and direction to NMSI provided curricular support materials | 44 |

Compliance with Site Coordinator stipend payment was determined by reviewing NMSI financial records, by school. One school did not request that a stipend be paid for the 2018–2019 school year; NMSI paid a Site Coordinator stipend to each of the other 47 program schools. For 85% of Program Schools, NMSI achieved the total seven points that could be earned from the school components of the fidelity matrix (see Table 47).

Table 47

*2018–2019 School Implementation Indicators*

| School components | Number of schools |
|---|---|
| Schools with a score of 7 | 41 |
| Schools with a score of 6 | 7 |

### b. Teacher Indicators: 2017–2018 Fidelity Matrix

Of the eight "teacher" matrix elements, four assess activities that directly influence classroom instruction and the remaining four evaluate NMSI's fulfillment of administrative responsibilities.

#### 1. Classroom Instruction

*Teacher Training:* One hundred and nineteen teachers attended all three training sessions, and 15 schools met the 80% threshold.

*Student Study Sessions:* In the survey, teachers were asked if they participated in the student study sessions. Fidelity at the school level was once again determined at the 80% threshold for survey respondents. Thirty-six schools were in compliance. Additionally, a later item in the survey asked participants how many student study sessions they had attended for each of the AP courses they teach. Ninety-seven teachers (or 53% of respondents to the item) said that they had attended three sessions for their primary AP course.

*Necessary Supplies:* Teachers were asked in the survey if they had the materials they needed to teach effectively. NMSI was considered compliant at a school if any teacher from the school said that they had the supplies they needed to teach their course. In 45 of 48 schools NMSI met these requirements, and 79% of survey participants who responded to this item said they had the materials needed to teach their course.

*NMSI Resource Availability:* There are several ways for teachers to obtain NMSI resources through the CRP:

- binders and handouts from training sessions
- work packets from the student study sessions
- shared drives (such as Google Drive) from facilitators and mentors
- websites curated by NMSI subject matter experts

In interviews, some teachers noted that much of the online content consists of the materials shared during training sessions. Therefore, we determined that lack of online engagement should

not be interpreted as a negative assessment of NMSI instructional resources as many teachers already had the materials they needed.

Survey participants were specifically asked if they were made aware of the online resources, and all schools are considered compliant with this matrix component. In subsequent survey items about how teachers incorporate the NMSI materials, 75% of respondents make use of the available materials in various manners.

### 2. NMSI Administration

*Access to Mentors:* Not all teachers wish to have a mentor provided by NMSI, so the mentoring component of the CRP is evaluated at the school level by determining if NMSI made a mentor available to any participating teacher, regardless of whether or not they decided they wanted one. Teacher survey participants in 45 of 48 schools said that mentoring was offered to them. Several teachers were matched to mentors in the three schools for which no survey respondent said that mentors were available. As a result, NMSI was considered in compliance for all 48 schools. However, 31% of respondents to the teacher survey item said that a mentor was not made available to them, perhaps highlighting an area for better outreach and communication around the availability of mentors.

*Payment of Stipends, Awards, and Bonuses:* NMSI's financial records indicate payments to teachers at each of the 48 schools for both the teacher stipend and the qualifying score award. Additionally, NMSI paid teacher bonuses at or above expectation for each of the 48 schools.

### 3. Overall Teacher Results

Four total points could be earned from each of the teacher component subgroups. NMSI was compliant in all 48 schools with the administrative components. Seventy-seven percent of schools were compliant with at least three of the classroom instruction components (see Table 48).

Table 48

*Classroom Instruction Components*

| Teacher-level component subtotal | Number of schools |
|---|---|
| Schools with a score of 4 | 12 |
| Schools with a score of 3 | 25 |
| Schools with a score of 2 | 10 |
| Schools with a score of 1 | 1 |

### c. Student Supports: 2018–2019 Fidelity Matrix

Of the four "student" fidelity matrix measures, one assesses student participation in student study sessions and the remaining three evaluate NMSI's fulfillment of administrative responsibilities.

#### 1. Student Participation

Data from NMSI's attendance tracking system tend to be inconsistent for the student study sessions. Through the survey process, however, schools had two opportunities to be considered in compliance with the category expectation.

In the student survey, participants were asked (a) how many student study sessions were offered for the course and (b) how many of the sessions they attended. For this metric, compliance was determined at the school level based on the number of students who self-reported attending all of the available sessions for the course. If 80% or more of the students surveyed attended all of the sessions, the school would have been in compliance. However, no schools met that threshold.

In the teacher and Site Coordinator surveys, participants were asked what percentage of students attended three student study sessions per course. The item was multiple choice with the highest range being "75%-100%." In 34 cases, a survey respondent selected "75%-100%" and the school was considered in compliance.

#### 2. NMSI Administration

The three administration functions that were considered student components of the fidelity matrix were the disbursement (as expected) of funds for classroom materials, exam fee subsidies, and student qualifying score awards. Funds were disbursed to each of the 47 schools that requested new materials in the 2018–2019 school year. As a result, all 48 schools were considered in compliance. Not all schools opt in to the exam fee subsidy component of the CRP. According to NMSI financial records, funds were disbursed to those schools expecting exam fee subsidies. All schools received funds from NMSI as expected for the qualifying scores achieved by their students, for compliance across all 48 schools.

#### 3. Overall Student Results

Seventy-one percent of schools achieved the four total points that could be earned from the student components of the matrix (see Table 49).

Table 49

*2018–2019 Student Implementation Indicators*

| Student components | Number of schools |
|---|---|
| Schools with a score of 4 | 34 |
| Schools with a score of 3 | 14 |

### d. Fidelity Matrix Summary Results

Considering all three component categories, 19 total points could be earned. As Table 50 illustrates, most schools (88%) achieved the matrix goal for 17 or more component targets.

Table 50

*2018–2019 Overall Fidelity Matrix Performance*

| Overall total | Number of schools | % of schools |
|---|---|---|
| Schools with a score of 19 | 10 | 21% |
| Schools with a score of 18 | 15 | 31% |
| Schools with a score of 17 | 17 | 35% |
| Schools with a score of 16 | 4 | 8% |
| Schools with a score of 15 | 2 | 4% |

## F. 2018–2019 Survey Results

In the 2018–2019 school year, to the best of our information, there were 324 teachers in the CRP in the schools included in this study. From this group, we received 211 completed teacher surveys.

### a. 2018–2019 Teacher Survey Response Summary

Sixty-five percent of teachers in Program Schools completed surveys with broad distribution among all ten states in the study (see Table 51). For survey participants, the 2018–2019 school year represented anywhere from their first to their third year in the program (see Figure 20).

Table 51

*Respondent Count by State*

| State | Teachers (*n*) | Surveys (*n*) | % of Teachers |
|---|---|---|---|
| California | 20 | 14 | 70% |
| Georgia | 27 | 18 | 67% |
| Illinois | 29 | 20 | 69% |
| Louisiana | 17 | 12 | 71% |
| Michigan | 27 | 19 | 70% |
| Missouri | 26 | 20 | 77% |
| North Dakota | 44 | 25 | 57% |
| Ohio | 32 | 24 | 75% |
| Pennsylvania | 46 | 35 | 76% |
| Texas | 56 | 24 | 43% |
| TOTAL | 324 | 211 | 65% |

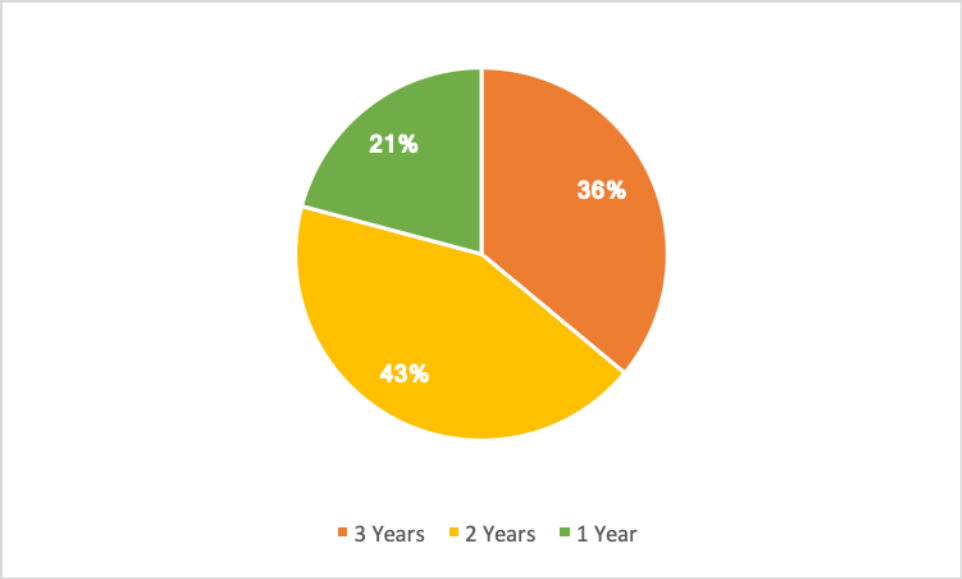



Figure 20. Distribution of survey participants by years participating in the CRP.

Survey participants were fairly evenly distributed across the three content areas included in the CRP (see Figure 21). Overall, the teacher respondents represent 9,872 AP students (see Table 52).

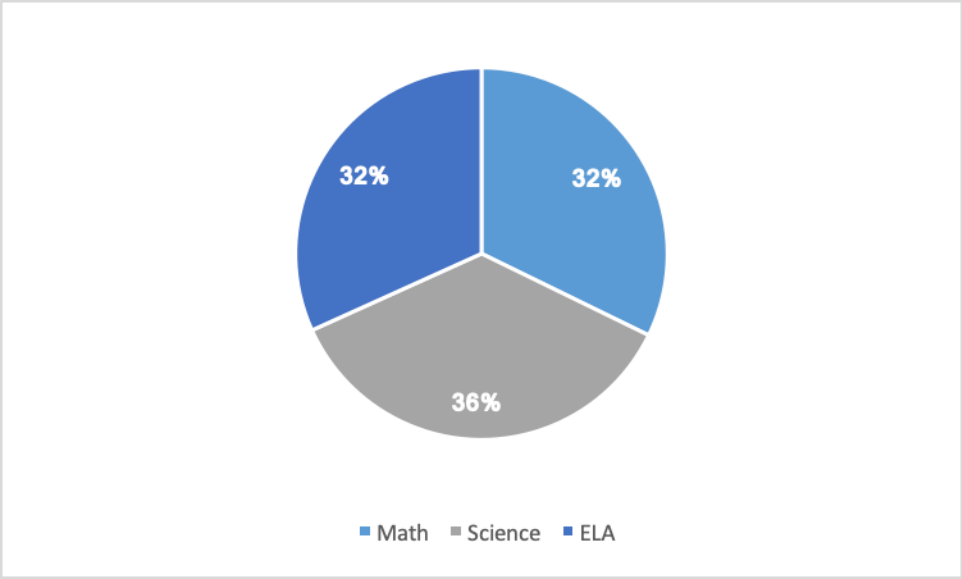



Figure 21. Distribution of survey participants by the content area in which they teach.

Table 52

*Grade-Level Distribution of Students Taught by 2019 Teacher Survey Participants*

| | 9th Grade | 10th Grade | 11th Grade | 12th Grade | Total |
|---|---|---|---|---|---|
| AP Students by Grade Level | 265 | 1,222 | 3,872 | 4,513 | 9,872 |

### 1. 2018–2019 Teacher Survey: Training

When asked to indicate in which CRP activities they participated, a fairly consistent percentage of respondents said they attended each of the three training sessions. The item did not control for the number of student study session the survey participants attended, so teachers typically had three opportunities to attend a session. In addition, attending the student study sessions often requires less travel than the three training sessions. As in previous years, participation in the mentor program fell far below participation in any of the other four activities (see Figure 22).

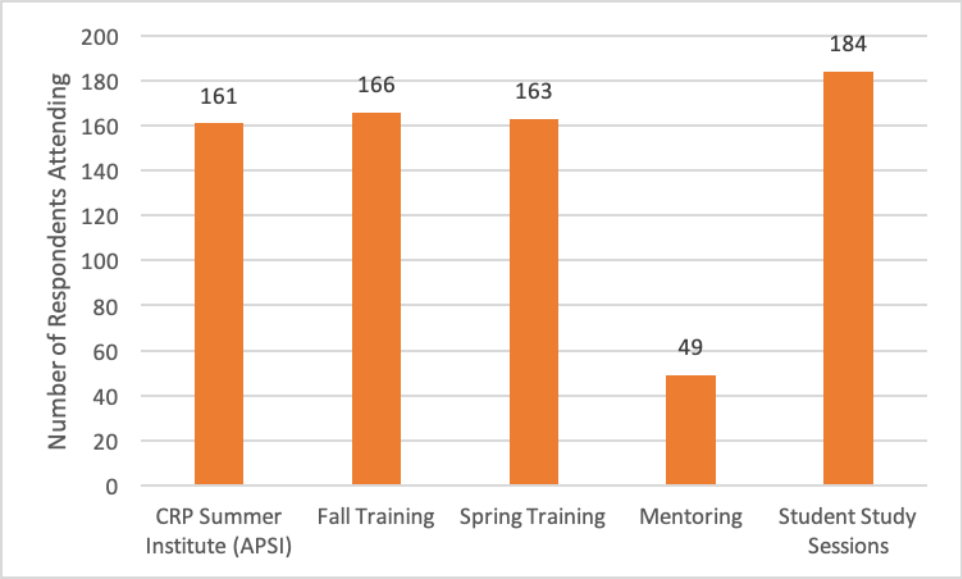



Figure 22. Self-reported attendance at CRP activities.

Figure 23 highlights both the drop off in Summer Institute attendance when comparing respondents with differing tenures in the CRP and the lower participation in student study sessions by first year CRP teachers, compared to other groups. Seventy percent of third year teachers participated in the summer training compared to 81% of second year teachers, and 77% of first year teachers. And while 96% of third year teachers reported attending the student study sessions, only 70% of first year teachers reported attending.

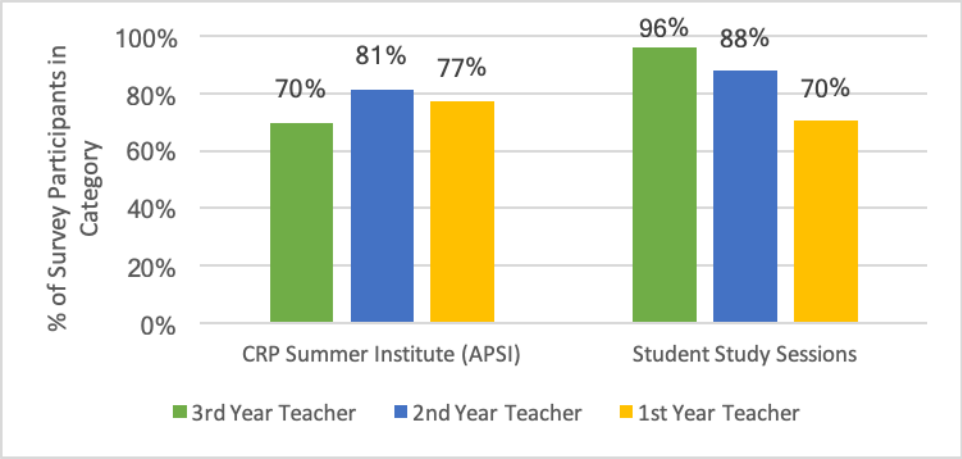



Figure 23. Self-reported attendance grouped by the teachers' year in the CRP.

Sixty-three third year teachers who took the survey for the 2018–2019 school year also completed the survey for the 2017–2018 school year. With the exception of the student study sessions, these participants reported attending activities at a lower rate in their third year. The

only activity for which the drop off was statistically significant was the summer institute (see Table 53).

Table 53

*Third-Year Teacher Activities 2018 and 2019 Surveys (*n = *63)*

| Activity | 2018 Survey *M (SD)* | 2019 Survey *M (SD)* | 2018-2019 diff | S.E. | t | df | *p* |
|---|---|---|---|---|---|---|---|
| Student Study Sessions | 0.95 (0.22) | 0.98 (0.13) | -0.032 | 0.022 | -1.426 | 62 | 0.159 |
| CRP Summer Institute | 0.91 (0.30) | 0.78 (0.42) | 0.127 | 0.042 | 3.003 | 62 | 0.004 |
| Fall Training | 0.89 (0.32) | 0.84 (0.37) | 0.048 | 0.048 | 1 | 62 | 0.321 |
| Spring Training | 0.83 (0.38) | 0.81 (0.40) | 0.015 | 0.048 | 0.331 | 62 | 0.742 |
| Mentoring | 0.32 (0.47) | 0.22 (0.42) | 0.095 | 0.059 | 1.624 | 62 | 0.109 |

Survey participants were asked to identify one CRP program element (in which they had previously indicated they had participated) as the most beneficial (see Figure 24). The summer training institute was selected as the most benefical program element by 108/199 responding teachers and was chosen by almost three times more teachers than the next most commonly-chosen program component (spring training).

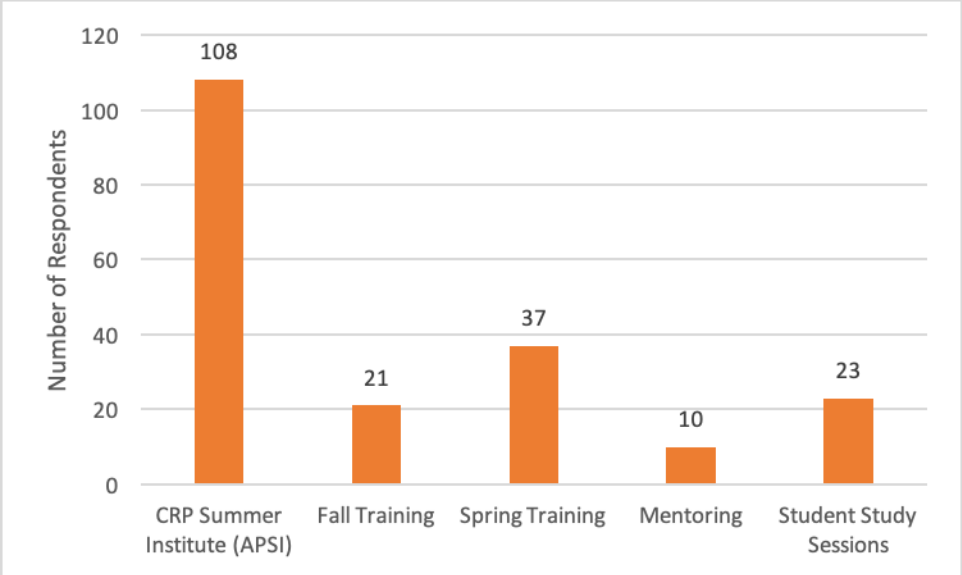



Figure 24. Single most beneficial CRP component (*n* = 199).

Survey participants were asked whether they agreed with a series of positive statements about each of the training sessions. The responses tabulated in Table 54 were limited to those who attended the given session, and percentages calculated based on attendance figures.

Table 54

*Teacher Opinions about the Training Sessions Attended*

| | Summer Session | | Fall Session | | Spring Session | |
|---|---|---|---|---|---|---|
| **2018–2019 Training Sessions** | ***n*** | **%** | ***n*** | **%** | ***n*** | **%** |
| Attended | 161 | | 166 | | 163 | |
| Scheduling & Location Convenient | 124 | 77% | 117 | 70% | 111 | 68% |
| Knowledgeable & Well-Prepared Facilitators | 148 | 92% | 138 | 83% | 142 | 87% |
| Improved My Content Knowledge | 132 | 82% | 116 | 70% | 108 | 66% |
| Felt More Qualified Afterwards | 136 | 84% | 120 | 72% | 117 | 72% |
| Clear Agenda & Goals | 144 | 89% | 137 | 83% | 142 | 87% |
| Effective Training Activities | 135 | 84% | 125 | 75% | 129 | 79% |
| Helped Me Differentiate Instruction | 109 | 68% | 89 | 54% | 87 | 53% |

There was a wide variation in the 2018–2019 data between teachers in their first, second, or third years in the program. As can be seen in Figure 25, however, one pattern that emerges is third year teachers were in general less satisfied with the fall session than they were with the other two sessions.

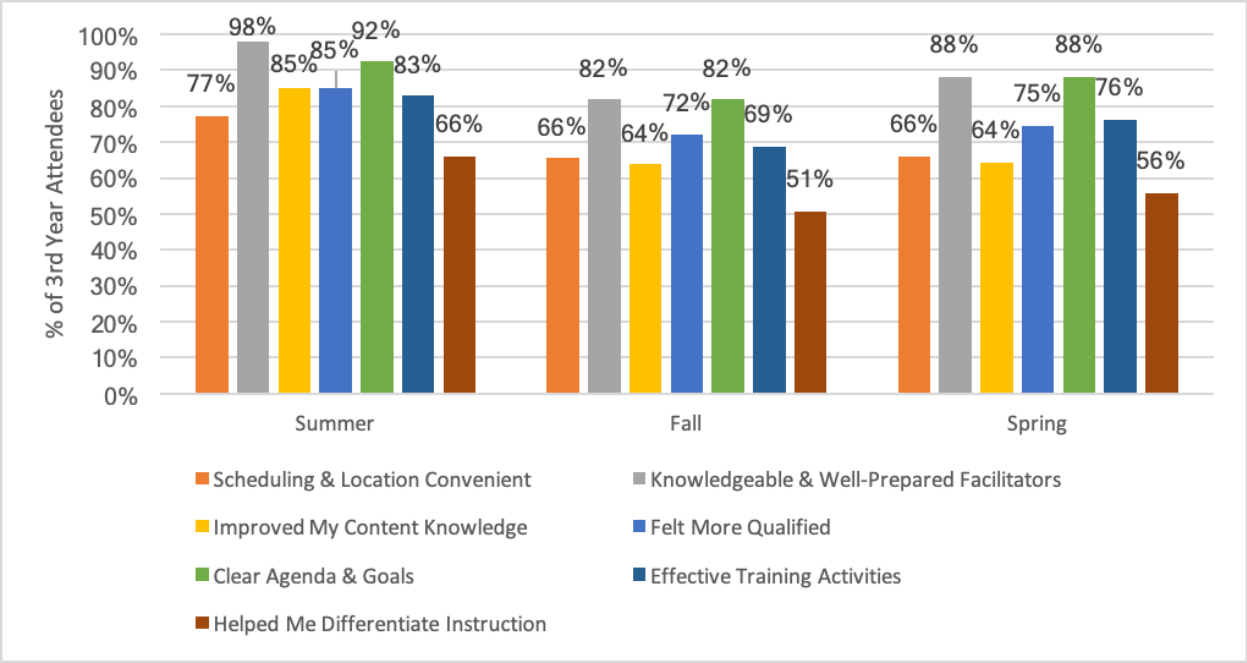



Figure 25. Third-year teacher evaluations of the three training sessions.

In addition, a significantly lower percentage of the third-year Treatment School teachers who completed both the 2018 and 2019 surveys selected the fall workshop as the most beneficial of the teacher activities in year three than selected it in year two (see Table 55).

Table 55

*Third-Year Teacher Most Beneficial Activity 2018 and 2019 Surveys (*n = *63)*

| **Label** | **2018 Survey *M (SD)*** | **2019 Survey *M (SD)*** | **2018-2019 diff** | **S.E.** | **t** | **df** | ***p*** |
|---|---|---|---|---|---|---|---|
| CRP Summer Institute | 0.56 (0.501) | 0.71(0.46) | -0.15 | 0.092 | -1.353 | 47 | 0.182 |
| Student Study Sessions | 0.19 (0.39) | 0.13 (0.34) | 0.057 | 0.056 | 1.524 | 58 | 0.133 |
| Fall Training | 0.18 (0.39) | 0.06 (0.23) | 0.125 | 0.056 | 2.203 | 48 | 0.032 |
| Spring Training | 0.14 (0.345) | 0.29 (0.46) | -0.159 | 0.086 | -1.734 | 46 | 0.09 |
| Mentoring | 0.05 (0.22) | 0.14 (0.36) | -0.093 | 0.1 | -1 | 9 | 0.343 |

By contrast, the same cohort of survey participants felt that the facilitators in the summer institute were knowledgeable at a significantly higher rate in 2019 than in 2018 (see Table 56).

Table 56

*Third-Year Teacher Summer Institute Knowledgeable Facilitators 2018 and 2019 Surveys (*n = *63)*

| **Training Session Evaluations: *The facilitators were knowledgeable and well-prepared*** | **2018 Survey *M (SD)*** | **2019 Survey *M (SD)*** | **2018-2019 diff** | **S.E.** | **t** | **df** | ***p*** |
|---|---|---|---|---|---|---|---|
| Four-Day AP Summer Institute (APSI) | 0.84 (0.37) | 0.98 (0.14) | -0.138 | 0.053 | -2.603 | 74.688 | 0.011 |
| Two-Day Fall Workshop | 0.88 (0.33) | 0.81 (0.40) | 0.064 | 0.07 | 0.907 | 101.984 | 0.367 |
| One-Day Spring Training | 0.85 (0.36) | 0.88 (0.325) | -0.036 | 0.068 | -0.532 | 100.134 | 0.596 |

Sixty-one of the teachers who completed the 2018–2019 survey also completed surveys in the the previous two study years. Categorical responses for this group of teachers were compared using a repeated measures ANOVA. Specifically, a linear mixed model with a covariate at Level-1 to indicate time points as well as a random effect term was fitted and an ANOVA was conducted to the fitted model. Results of Tukey post-hoc tests are reported. For two of the training sessions (fall and spring), teacher responses suggest a perception of diminishing returns (see Table 57). For the summer institute, however, third year teachers felt that the facilitators were knowledgeable and well prepared at a higher rate in the summer of 2018 than in the summer of 2017 (see Table 58 and Figure 26).

Table 57

*ANOVA—Opinions about Training Sessions (*n = *61)*

| Training session | 2016–2017 | | 2017–2018 | | 2018–2019 | | Chi (df) | Significant pairwise comparisons |
|---|---|---|---|---|---|---|---|---|
| | ***n*** | **%** | ***n*** | **%** | ***n*** | **%** | | |
| The facilitators were knowledgeable and well-prepared | | | | | | | | |
| Four-Day AP Summer Institute (APSI) | 50 | 90.9 | 47 | 85.5 | 46 | 97.9 | 8.7 (2)* | yr3 > yr2 |
| Two-Day Fall Workshop | 50 | 89.3 | 47 | 87.0 | 41 | 80.4 | 2 (2) | |
| One-Day Spring Training | 53 | 93.0 | 42 | 84.0 | 43 | 87.8 | 2.9 (2) | |
| The training activities improved my content knowledge | | | | | | | | |
| Four-Day AP Summer Institute (APSI) | 48 | 87.3 | 40 | 72.7 | 41 | 87.2 | 5.5 (2) | |
| Two-Day Fall Workshop | 46 | 82.1 | 41 | 75.9 | 31 | 60.8 | 7.3 (2)* | yr3 < yr1; yr3 < yr2 |
| One-Day Spring Training | 42 | 73.7 | 34 | 68.0 | 30 | 61.2 | 2.4 (2) | |
| At the end of the training, I felt more qualified to be an effective AP instructor | | | | | | | | |
| Four-Day AP Summer Institute (APSI) | 48 | 87.3 | 41 | 74.5 | 41 | 87.2 | 5.1 (2) | |
| Two-Day Fall Workshop | 47 | 83.9 | 39 | 72.2 | 35 | 68.6 | 4.8 (2) | |
| One-Day Spring Training | 48 | 84.2 | 32 | 64.0 | 35 | 71.4 | 6.5 (2) | yr2 < yr1 |
| The training activities were effective | | | | | | | | |
| Four-Day AP Summer Institute (APSI) | 46 | 83.6 | 41 | 74.5 | 39 | 83.0 | 2.1 (2) | |
| Two-Day Fall Workshop | 47 | 83.9 | 38 | 70.4 | 33 | 64.7 | 7.2 (2)* | yr3 < yr1 |
| One-Day Spring Training | 43 | 75.4 | 35 | 70.0 | 36 | 73.5 | 0.4 (2) | |

Table 58

*Pairwise Comparison—Summer Institute Facilitators Knowledgeable*

| contrast 2 | Estimate | S.E. | df | T | *p* |
|---|---|---|---|---|---|
| yr2-yr1 | -0.06 | 0.05 | 102 | -1.13 | 0.26 |
| yr3-yr1 | 0.07 | 0.05 | 107 | 1.27 | 0.21 |
| yr3-yr2 | 0.12 | 0.05 | 100 | 2.37 | 0.02 |

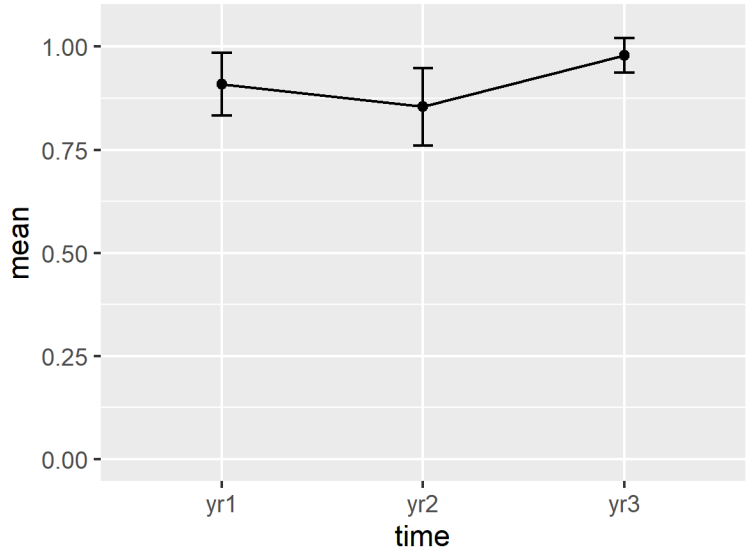



Figure 26. Change across three years—Summer Institute facilitators knowledgeable.

For the fall training, respondents in the final study year felt that the training activities were effective and that the training improved their content knowledge at significantly *lower* rates than in previous years (see Table 59, Table 60, Figure 27, and Figure 28). This may be in part because as teachers continued to participate in the program, they had less growth than they had initially and so the activities were not seem to improve content knowledge as much. Perhaps because teachers felt their content knowledge had increased a lot the first year, but at lower rates in subsequent years because their new baseline was higher.

Table 59

*Pairwise Comparison—Fall Workshop Improved My Content Knowledge*

| **contrast 10** | **Estimate** | **S.E.** | **df** | **T** | ***p*** |
|---|---|---|---|---|---|
| yr2-yr1 | -0.06 | 0.08 | 100 | -0.80 | 0.42 |
| yr3-yr1 | -0.22 | 0.08 | 103 | -2.77 | 0.01 |
| yr3-yr2 | -0.16 | 0.08 | 100 | -1.97 | 0.05 |

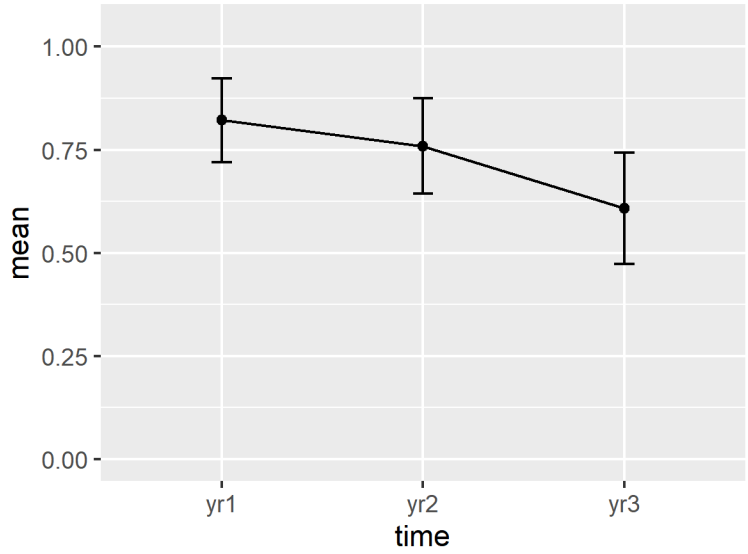



Figure 27. Change across three years—fall training improved my content knowledge.

Table 60

*Pairwise Comparison—Fall Training Activities Were Effective*

| **contrast 13** | **Estimate** | **S.E.** | **df** | **t** | ***p*** |
|---|---|---|---|---|---|
| yr2-yr1 | -0.14 | 0.07 | 101 | -1.88 | 0.06 |
| yr3-yr1 | -0.20 | 0.07 | 104 | -2.62 | 0.01 |
| yr3-yr2 | -0.06 | 0.08 | 101 | -0.78 | 0.44 |

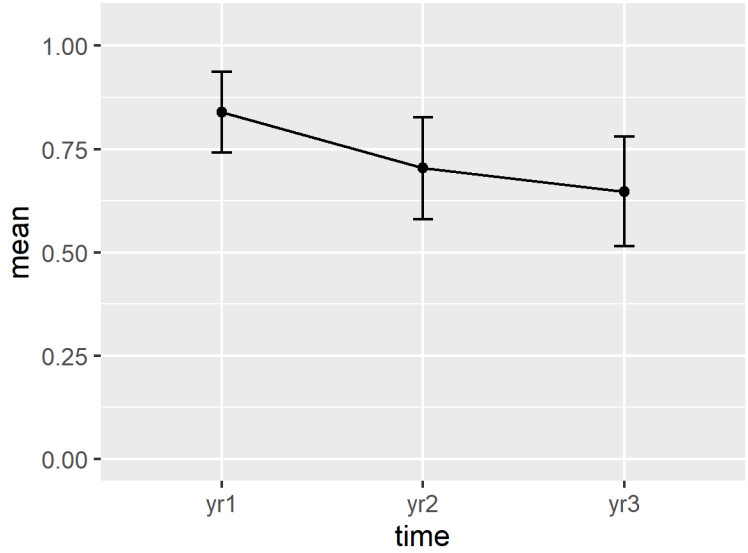



Figure 28. Change across three years—fall workshop activities were effective.

For the spring training, survey participants felt more qualified to be effective AP instructors at a significantly lower rate in year two than in year one. While this figure crept up in Year 3, the change was not significant when compared to Year 2 or Year 1 (see Table 61 and Figure 29).

Table 61

*Pairwise Comparison—After Spring Training Teacher Felt More Qualified*

| **contrast 18** | **estimate** | **S.E.** | **df** | **t** | ***p*** |
|---|---|---|---|---|---|
| yr2-yr1 | -0.20 | 0.08 | 104 | -2.50 | 0.01 |
| yr3-yr1 | -0.13 | 0.08 | 105 | -1.64 | 0.10 |
| yr3-yr2 | 0.07 | 0.08 | 102 | 0.81 | 0.42 |

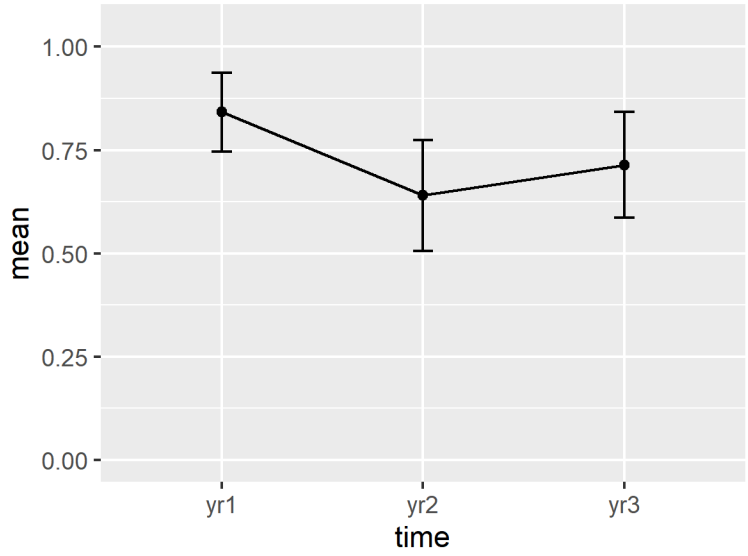



Figure 29. Change across three years—after spring training teacher felt more qualified.

Survey participants were asked in what ways training activities had changed from the 2017–2018 school year. In addition to the five given options, respondents could select no changes, do not know, or other—with the option to elaborate on the "other" designation (see Figure 30). Respondents could also select more than one option, and responses were only tabulated for teachers in their third or second year of the program (as they had a prior year for comparison). The participants who selected "no changes" represented teachers in all content areas. Most commonly noted changes from the 2017–2018 school year were the addition of new or different topics.

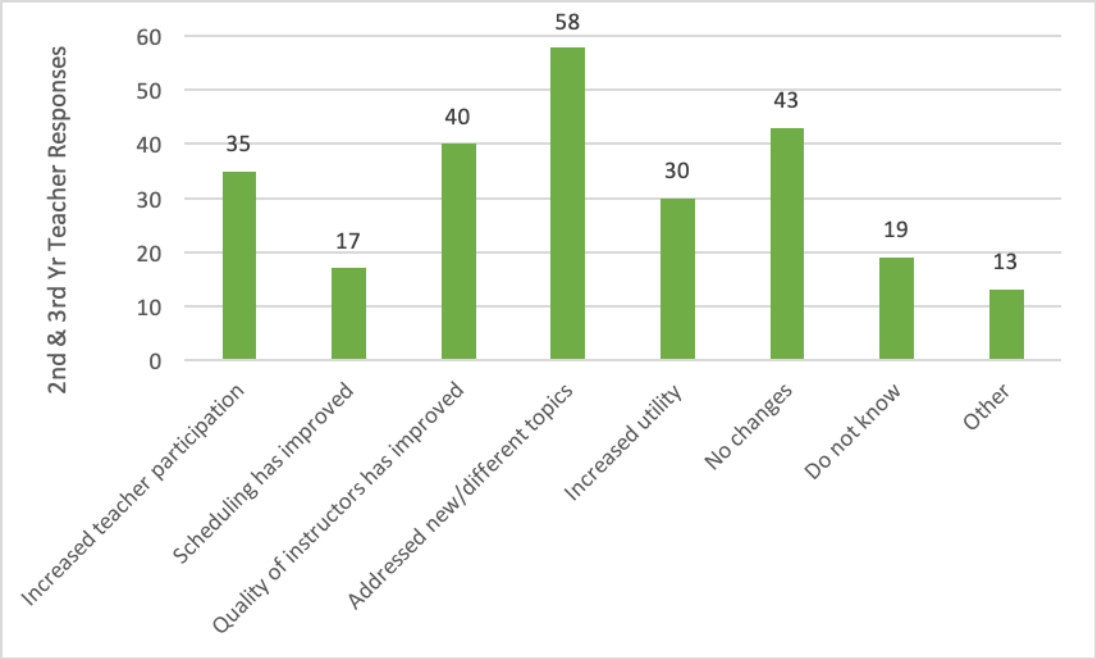



Figure 30. Changes to training sessions from the 2017–2018 school year.

Survey participants were given the opportunity to comment on how training had changed from the 2017–2018 school year, and opinions varied (see Appendix C). However, English instructors' comments were fairly consistently negative about the training received.

### 2. 2018–2019 Teacher Survey: Mentoring

The series of mentoring items first asked survey participants to think of how (and how frequently) they would like to communicate with a mentor. The two subsequent items inquired as to whether the respondent had the option to participate in mentoring and if the CRP Program Manager assisted with matching the teacher to a mentor. Survey respondents were then asked if their preferred forms of communication were available through the mentoring program (regardless of the preferred communication method). The majority of respondents to the communication method item (84%) said their preferred method of communication was available (see Table 62).

Table 62

*Information About the CRP Mentoring Program*

| Mentoring | Yes | | No | | Blank | |
|---|---|---|---|---|---|---|
| | # | % | # | % | # | % |
| Was mentoring available? | 134 | 64% | 59 | 28% | 18 | 9% |
| Did PM help match to mentor? | 74 | 35% | 64 | 30% | 73 | 35% |
| Was preferred communication available? | 105 | 50% | 20 | 9% | 86 | 41% |

Across the board, participants responding to items about the convenience and frequency of mentor contact rated the program fairly lowly (see Table 63).

Table 63

*Frequency and Convenience of Mentor Meetings*

| Mentor program evaluation | *n* | *M* | *SD* |
|---|---|---|---|
| My mentor made sufficient time to meet with me, and respond to questions | 100 | 2.97 | 1.06 |
| The mentoring sessions were conveniently scheduled | 102 | 2.93 | 1.03 |
| I met with my mentor as frequently as I wanted to | 101 | 2.91 | 1.04 |
| If mentor meetings were in-person, I would have met more frequently with my mentor | 100 | 2.53 | 1.13 |

*Note.* 1 = *strongly disagree*, 2 = *disagree somewhat*, 3 = *agree somewhat*, and 4 = *strongly agree*.

Most survey participants also reported a low frequency of contact with their mentors (see Figure 31).

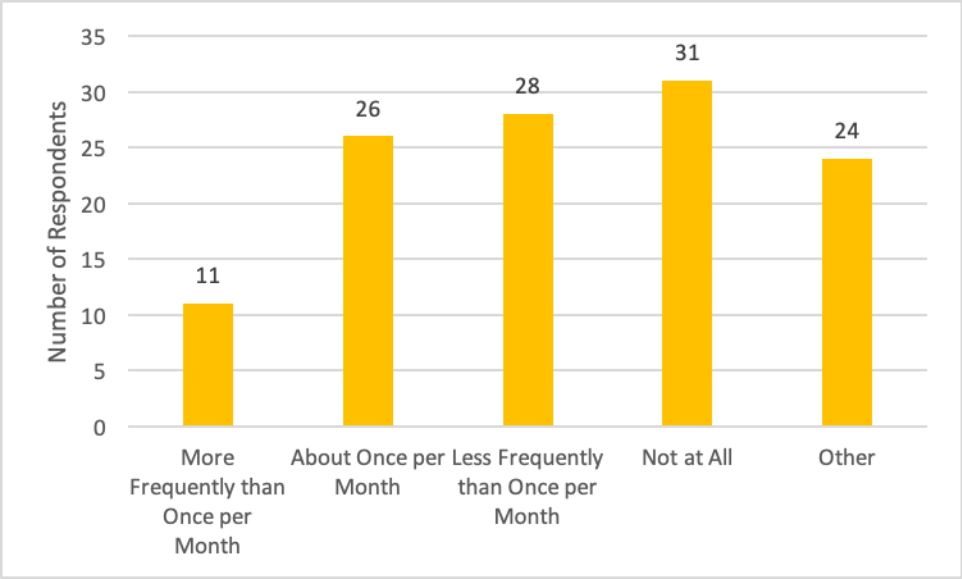



Figure 31. Frequency with which survey participants communicated with their mentors.

Despite the apparent infrequent contact with mentors, participants rated the mentor program fairly highly; the minimum average rating of all mentoring statements was 3.00 or *agree somewhat* (see Table 64).

Table 64

*Participant Evaluations of the Mentoring Program*

| Mentor Evaluation | *n* | *M* | *SD* |
|---|---|---|---|
| The mentor was knowledgeable and well-prepared | 80 | 3.56 | 0.72 |
| The mentor's guidance on pacing was helpful | 79 | 3.24 | 0.93 |
| The mentoring improved my content knowledge | 79 | 3.13 | 0.91 |
| The mentoring honed my instructional skills and techniques | 80 | 3.03 | 0.99 |
| Because of the mentoring I am a more effective and qualified AP instructor | 80 | 3.00 | 1.02 |

*Note.* 1 = *strongly disagree*, 2 = *disagree somewhat*, 3 = *agree somewhat*, and 4 = *strongly agree*.

### 3. 2018–2019 Teacher Survey: Student Study Sessions

Survey participants were asked how AP courses they were teaching and then how many student study sessions they attended for each AP course. Fourteen percent of respondents to the item ($n$ = 184) reported teaching more than one CRP AP course in the 2018–2019 school year (see Table 65).

Table 65

*Number of CRP AP Courses Taught by Content Area*

| Content area | 1 AP course | 2 AP courses | 3 or more AP courses |
|---|---|---|---|
| ELA | 54 | 4 | 0 |
| Science | 61 | 4 | 3 |
| Math | 43 | 14 | 1 |
| TOTAL | 158 | 22 | 4 |

Survey participants attended all three of the student study sessions offered for 52% of the 214 CRP-related AP courses they taught, and reported attending zero sessions for only 19 of the courses taught (9%). See Figure 32.

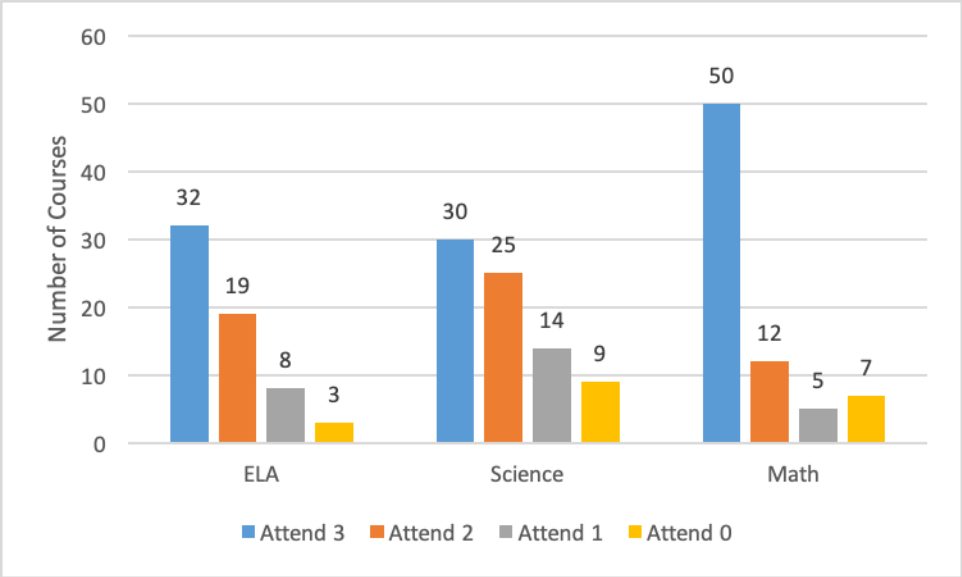



Figure 32. Number of student study sessions attended by teachers in each content area.

Students are expected to attend all three student study sessions for each CRP-supported AP course. Survey respondents ($n$ = 204) were asked to select the range that represented the percentage of their students who managed to attend all three sessions. Thirty-four percent of participants said that fewer than 25% of their students attended all three student study sessions (see Figure 33).

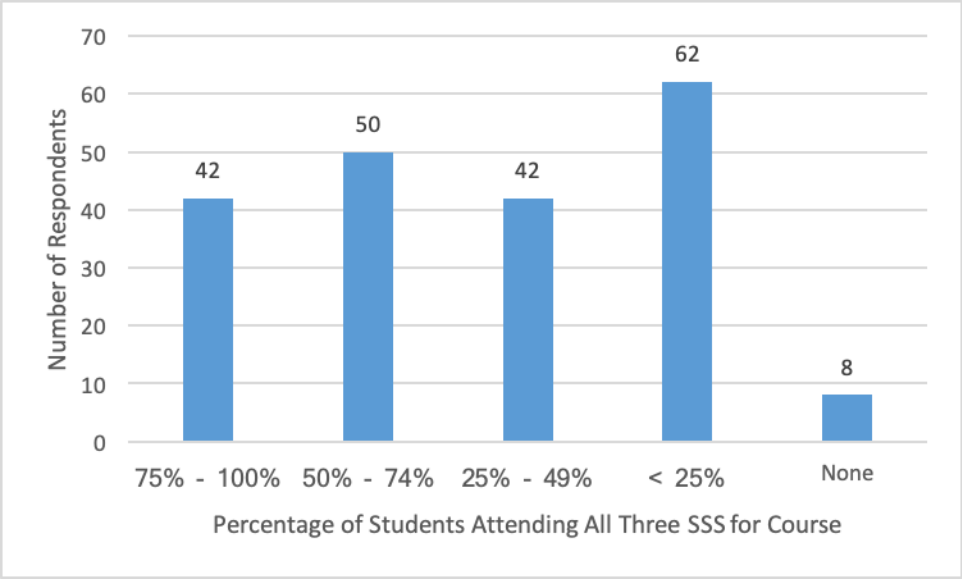



Figure 33. Percentage range of students who attended all three sessions.

In fact the 61 Treatment School teachers who took the survey all three years noted a significant decline in the percentage of students who attended all three student study sessions from Year 1 to Year 3 (see Table 66, Table 67, and Figure 34).

Table 66

*ANOVA—Student Attendance at Student Study Sessions*

| | 2016–2017 | | 2017–2018 | | 2018–2019 | | | |
|---|---|---|---|---|---|---|---|---|
| **Student attendance** | ***M*** | ***SD*** | ***M*** | ***SD*** | ***M*** | ***SD*** | **F (df)** | **Significant pairwise comparisons** |
| Attend 3 Sessions | 2.7 | 1.1 | 2.6 | 1.1 | 2.3 | 1.1 | 4.2 ( 2 ) | Yr 3 < yr1 |

*Note*. 0 = None, 1 = <25%, 2 = 25%-49%, 3 = 50%-74%, 4 = 75%-100%.

Table 67

*Pairwise Comparison—Student Attendance at Student Study Sessions*

| **contrast** | **estimate** | **S.E.** | **Df** | **t** | ***p*** |
|---|---|---|---|---|---|
| yr2-yr1 | -0.16 | 0.14 | 119 | -1.13 | 0.26 |
| yr3-yr1 | -0.41 | 0.14 | 119 | -2.87 | 0.00 |
| yr3-yr2 | -0.25 | 0.14 | 119 | -1.74 | 0.08 |

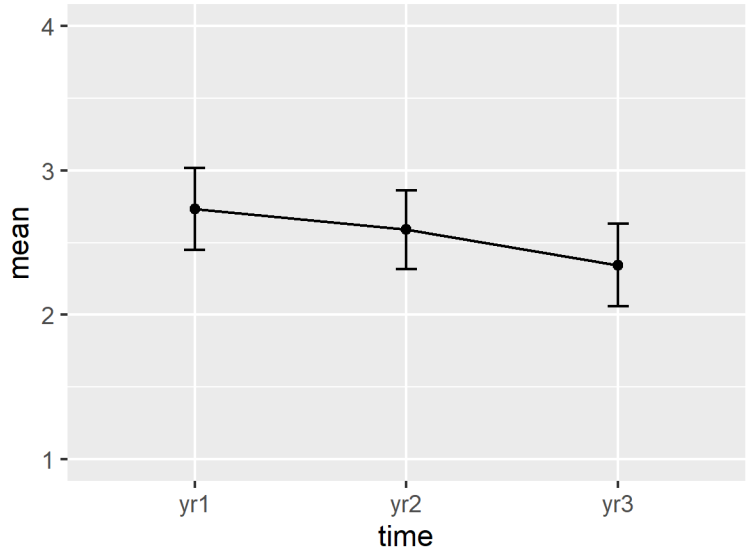



Figure 34. Change across three years—student attendance at student study sessions.

Survey participants were asked whether they agreed with a series of positive qualitative statements about the student study sessions. The three most positive assessments highlight the benefits of the sessions for students (see Table 68). The results were fairly unchanged from the 2018 survey. The lone exception was that the mean response to the statement about the convenience of the sessions dropped from 2.90 (*SD* = 0.86) to 2.68 in the 2018–2019 school year.

Table 68

*Evaluations of Student Study Sessions*

| Teacher evaluations of student study sessions | *n* | *M* | *SD* |
|---|---|---|---|
| The study sessions were led by AP experts who taught NMSI-created lessons | 195 | 3.48 | 0.67 |
| The study sessions improved students' content knowledge | 196 | 3.27 | 0.67 |
| The study sessions helped to increase student confidence | 194 | 3.23 | 0.73 |
| The study sessions highlighted the instructional needs of the students for me to continue addressing in class | 194 | 3.13 | 0.81 |
| Students were active participants (e.g., answering and asking questions, focused on tasks assigned, etc.) | 194 | 3.07 | 0.83 |
| I was able to take the strategies I saw employed during the study sessions back to my own classrooms to improve student achievement | 192 | 3.01 | 0.87 |
| I learned a great deal from watching the expert teachers during the student study sessions | 192 | 2.90 | 0.94 |
| The study sessions were conveniently scheduled to accommodate student schedules | 197 | 2.68 | 0.90 |
| My input was considered when determining the study session topics | 195 | 1.79 | 0.97 |

*Note.* 1 = *strongly disagree*, 2 = *disagree somewhat*, 3 = *agree somewhat*, and 4 = *strongly agree*.

When we compare responses of third year teachers at Treatment Schools ($n$ = 76) with second year teachers at Treatment Schools ($n$ = 32), we see that the teachers with more years in the program feel more strongly that the student study sessions highlight the instructional needs of their students (see Table 69).

Table 69

*Treatment Schools: Evaluation of Sessions by Participant Year in Program*

| Student study session evaluation | 2nd Year *M (SD)* | 3rd Year *M (SD)* | 2nd Yr – 3rd Yr diff | S.E. | t | Df | *p* |
|---|---|---|---|---|---|---|---|
| The study sessions highlighted the instructional needs of the students for me to continue addressing in class | 3 (0.845) | 3.361 (0.657) | -0.361 | 0.175 | -2.064 | 42.284 | 0.045 |

*Note.* 1 = *strongly disagree*, 2 = *disagree somewhat*, 3 = *agree somewhat*, and 4 = *strongly agree*.

To further the supposition that opinions about the student study sessions improve with experience, the student study session ratings of the 61 Treatment School teachers who completed the survey all three years significantly increased in Years 2 and 3 compared to the first year on three points: quality of instructors, positive impact on students' content knowledge, and highlighting student needs. All of this despite feeling that the scheduling of the sessions had only

gotten worse since the second year (see Table 70 through Table 74 and Figure 35 through Figure 38).

Table 70

*ANOVA—Opinions About Student Study Sessions*

| Statement | 2016–2017 | | | 2017–2018 | | | 2018–2019 | | | F (df) | Significant pairwise comparisons |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | ***n*** | ***M*** | ***SD*** | ***n*** | ***M*** | ***Sd*** | ***n*** | ***M*** | ***Sd*** | | |
| The study sessions were led by AP experts who taught NMSI-created lessons | 60 | 3.2 | 1.0 | 61 | 3.6 | 0.5 | 59 | 3.6 | 0.5 | 5.4 (2) | yr2>yr1; yr3>yr1 |
| The study sessions improved students' content knowledge | 60 | 3.1 | 0.9 | 61 | 3.4 | 0.6 | 60 | 3.4 | 0.6 | 3.8 (2) | yr2>yr1; yr3>yr1 |
| The study sessions highlighted the instructional needs of the students for me to continue addressing in class | 60 | 2.9 | 0.9 | 60 | 3.2 | 0.6 | 59 | 3.4 | 0.7 | 6.2 (2) | yr2>yr1; yr3>yr1 |
| The study sessions were conveniently scheduled to accommodate student schedules | 60 | 2.8 | 0.9 | 61 | 3.0 | 0.8 | 60 | 2.6 | 0.9 | 3.9 (2) | yr3<yr2 |

*Note.* 1 = *strongly disagree*, 2 = *disagree somewhat*, 3 = *agree somewhat*, and 4 = *strongly agree*.

Table 71

*Pairwise Comparison—Student Study Sessions Led by AP Experts*

| contrast | Estimate | S.E. | df | t | *p* |
|---|---|---|---|---|---|
| yr2-yr1 | 0.37 | 0.13 | 117 | 2.91 | 0.00 |
| yr3-yr1 | 0.36 | 0.13 | 119 | 2.76 | 0.01 |
| yr3-yr2 | -0.02 | 0.13 | 118 | -0.13 | 0.90 |

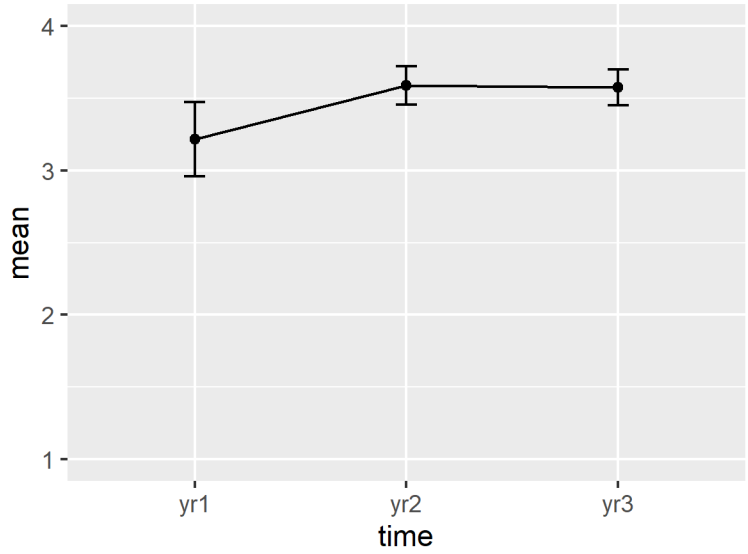



Figure 35. Change across three years—student study sessions led by AP experts.

Table 72

*Pairwise Comparison—Student Study Sessions Improved Student Content Knowledge*

| contrast | estimate | S.E. | df | t | *p* |
|---|---|---|---|---|---|
| yr2-yr1 | 0.29 | 0.12 | 118 | 2.46 | 0.02 |
| yr3-yr1 | 0.28 | 0.12 | 119 | 2.32 | 0.02 |
| yr3-yr2 | -0.01 | 0.12 | 118 | -0.12 | 0.90 |

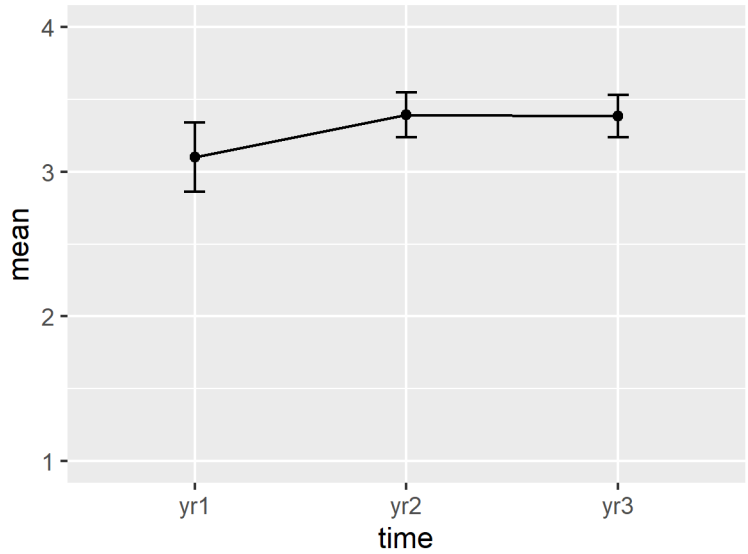



Figure 36. Change across three years—student study sessions improved student content knowledge.

Table 73

*Pairwise Comparison—Student Study Sessions Highlighted Student Needs*

| contrast | Estimate | S.E. | df | t | p |
|---|---|---|---|---|---|
| yr2-yr1 | 0.30 | 0.12 | 117 | 2.51 | 0.01 |
| yr3-yr1 | 0.41 | 0.12 | 117 | 3.40 | 0.00 |
| yr3-yr2 | 0.11 | 0.12 | 117 | 0.91 | 0.37 |

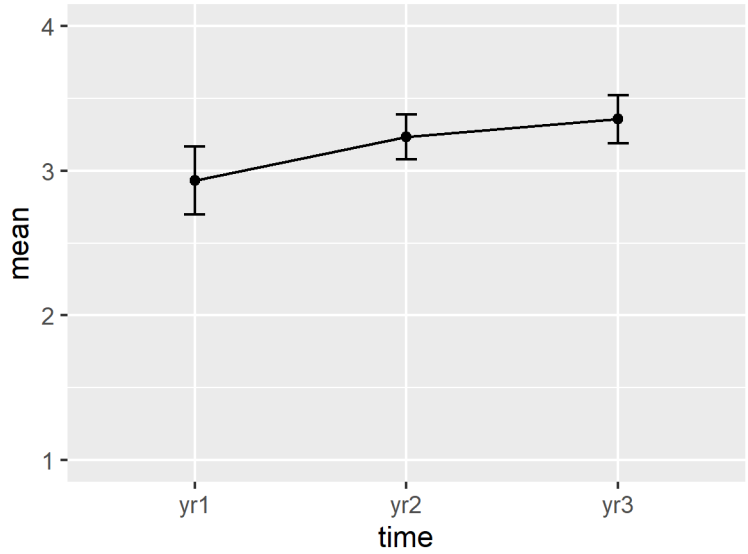



Figure 37. Change across three years—student study sessions highlighted student needs.

Table 74

*Pairwise Comparison—Student Study Sessions Conveniently Scheduled*

| contrast | Estimate | S.E. | df | t | *p* |
|---|---|---|---|---|---|
| yr2-yr1 | 0.14 | 0.13 | 118 | 1.12 | 0.26 |
| yr3-yr1 | -0.21 | 0.13 | 119 | -1.65 | 0.10 |
| yr3-yr2 | -0.36 | 0.13 | 118 | -2.78 | 0.01 |

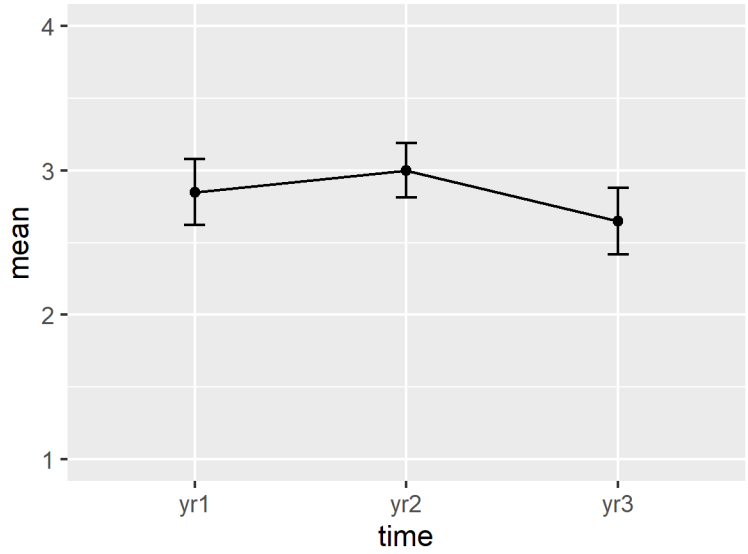



Figure 38. Change across three years—student study sessions conveniently scheduled.

Respondents were asked to assess the usefulness of the student study sessions, both for the students and for themselves. Survey participants confirmed that they found the sessions more beneficial for the students (see Figure 39).

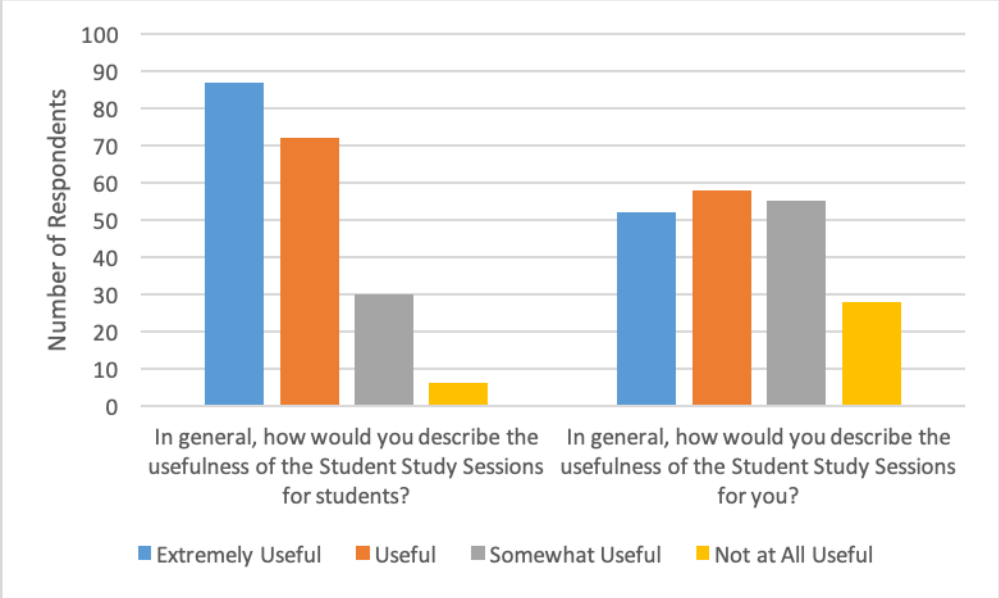



Figure 39. Degree to which the student study sessions were useful.

In a multiple choice, multiple selection item, survey participants were asked to indicate how the student study sessions had changed from the 2017–2018 school year. Sixty-eight second

and third year teachers said that the sessions had not changed year over year; the next two most frequently selected choices were fewer or more students attending the sessions (see Figure 40).

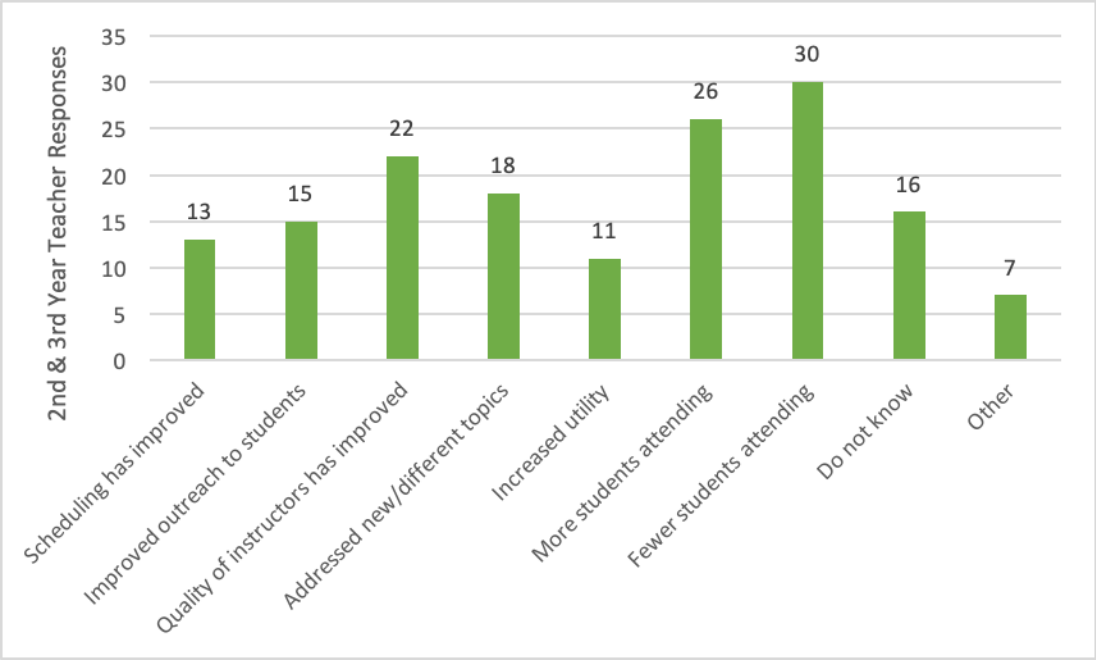



Figure 40. Changes to student study sessions from the 2017–2018 school year.

A significantly lower percentage of the 63 third year Treatment School teachers who completed both the 2018 and 2019 surveys observed changes to session topics in Year 3 than in Year 2 (see Table 75).

Table 75

*Third-Year Teacher Student Study Sessions Changes 2018 and 2019 Surveys (*n = *63)*

| **Changes to student study sessions** | **2018 Survey *M (SD)*** | **2019 Survey *M (SD)*** | **2018-2019 diff** | **S.E.** | **t** | **df** | ***p*** |
|---|---|---|---|---|---|---|---|
| Addressed new/different topics | 0.25 (0.44) | 0.13 (0.34) | 0.123 | 0.06 | 2.206 | 59 | 0.031 |

## 4. 2018–2019 Teacher Survey: Instructional Materials and Equipment

Survey participants were asked a series of questions about their experience with NMSI funding for classroom and lab materials and equipment. Of the 210 participants who responded to an item asking if they had received any materials through the CRP, 72% said yes. In the 2017–2018 survey, 80% of respondents had received materials. Participants were asked to indicate if

they had obtained equipment or materials in a limited list of categories during the school year with NMSI funding (see Figure 41). Two respondents provided additional comments:

- Novels, dry erase boards
- The books that I requested were denied. So they sent me some random titles that I will try to utilize. I am extremely disappointed that I was not given more freedom in my requests. This has always been a huge part of the appeal of NMSI, and to not be trusted with my choices, to not have my required texts fulfilled felt like a slap in the face.

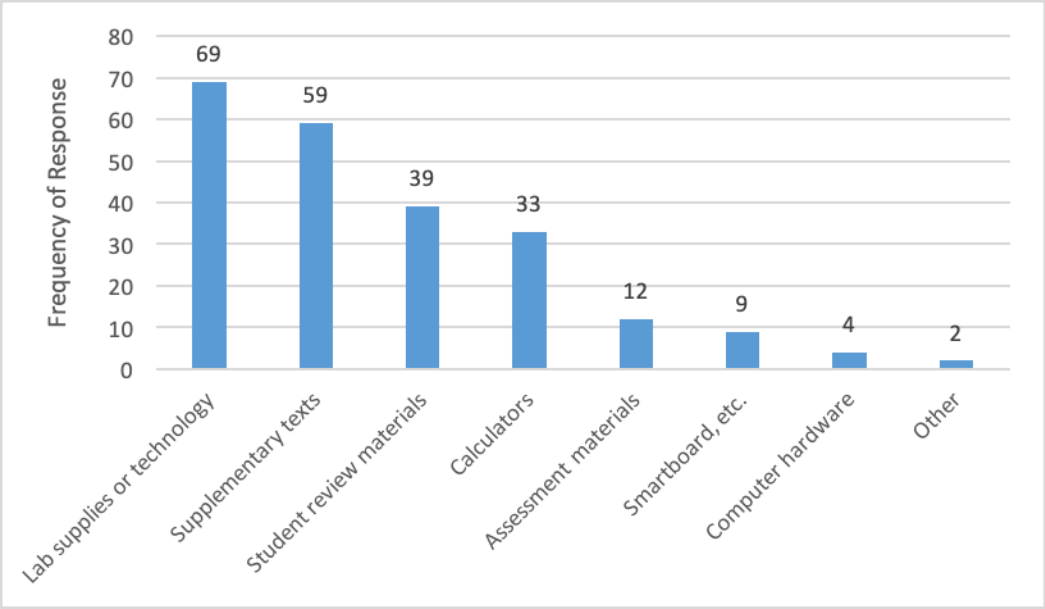



Figure 41. Materials and equipment obtained with NMSI financial resources.

Survey participants were also asked if they have the materials necessary to teach their course. Of the 198 teachers responding to the item, 79% said yes.

### 5. 2018–2019 Teacher Survey: Financial Stipends and Awards

Respondents were asked to rate how important the CRP incentives were in encouraging them to teach the AP course. Consistent with feedback in teacher interviews, 79% of survey participants said that the incentives were only somewhat important or not at all important (see Figure 42). However, interviewees tended to stress that the incentives were a very positive component of the program even if they did not weigh heavily on their decision whether or not to teach an AP course.

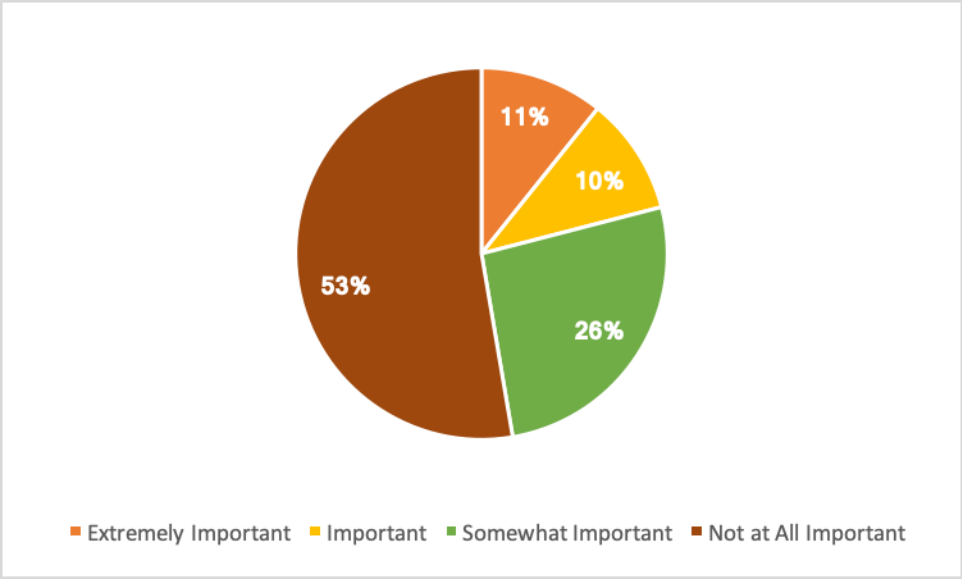



Figure 42. Importance of CRP financial incentives in teachers' decision to teach the AP course.

In fact, for the 61 Treatment School teachers who completed the survey each year the importance of the incentives decreased significantly from the first year of the program (see Table 76, Table 77, and Figure 43).

Table 76

*ANOVA—Importance of Financial Incentives*

| | 2016–2017 | | | 2017–2018 | | | 2018–2019 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Statement** | ***n*** | ***M*** | ***SD*** | ***n*** | ***M*** | ***SD*** | ***n*** | ***M*** | ***SD*** | **F (df)** | **Significant pairwise comparisons** |
| Importance of incentives | 52 | 2.0 | 0.9 | 61 | 2.0 | 1.0 | 60 | 1.8 | 1.0 | 4.1 (2)* | yr2 < yr1; yr3 < yr1 |

*Note.* 1 = *Not at all Important*, 2 = *Somewhat Important*, 3 = *Important*, 4 = *Extremely Important*.

Table 77

*Pairwise Comparison—Importance of Financial Incentives*

| **contrast** | **Estimate** | **S.E.** | **df** | **t** | ***p*** |
|---|---|---|---|---|---|
| yr2-yr1 | -0.08 | 0.11 | 112 | -0.74 | 0.46 |
| yr3-yr1 | -0.31 | 0.11 | 111 | -2.71 | 0.01 |
| yr3-yr2 | -0.22 | 0.11 | 110 | -2.08 | 0.04 |

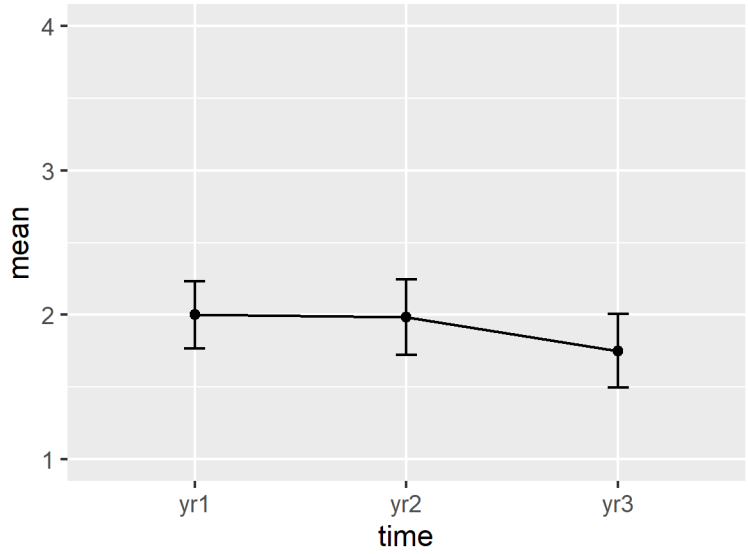



Figure 43. Change across three years—importance of financial incentives.

Survey participants were asked how financial awards had changed year over year. Very few second and third year teachers selected multiple choice options other than "no change" and "do not know". In addition to the multiple choice selections, respondents had the option to comment (see Appendix D).

In many school districts, students are eligible for exam fee subsidies through the CRP or other entities. Of the 122 respondents who said that students in their school were eligible for subsidies from NMSI, 60 also said that students could get subsidies through their school, district, or other source. A total of 71 respondents who did not believe that students received exam fee subsidies through the CRP said that students received subsidies from alternative sources. Teachers in every school ($n$ = 193) said that students were eligible for an exam fee subsidy from at least one source.

### 6. 2018–2019 Teacher Survey: Use of CRP Instructional Resources

There are many ways for teachers to obtain instructional materials through the CRP such as in the binders from the summer sessions, on the program's website, in instructional packets for student study sessions, or through Google drives maintained by mentors and facilitators. In some cases, survey items were designed to determine survey participants use of a specific tool (the NMSI teacher website) while other items were designed to evaluate respondents use of resources regardless of how they were obtained.

Ninety-two percent of respondents said that they were given access to online resources (the NMSI teacher website) through the CRP. Of these 195 survey participants, 38% accessed CRP online resources more frequently than once per month (see Figure 44).

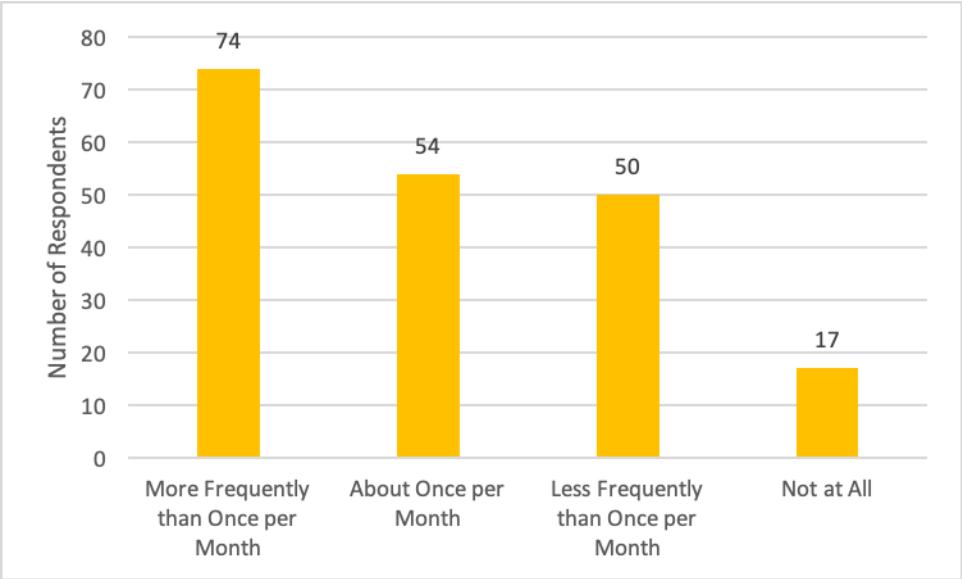



Figure 44. Frequency with which survey participants access the CRP teacher website.

Survey participants were asked in a multiple-choice, multiple-selection item how they incorporated NMSI resources into instruction. Adding depth to instruction was on par with exam preparation (see Figure 45).

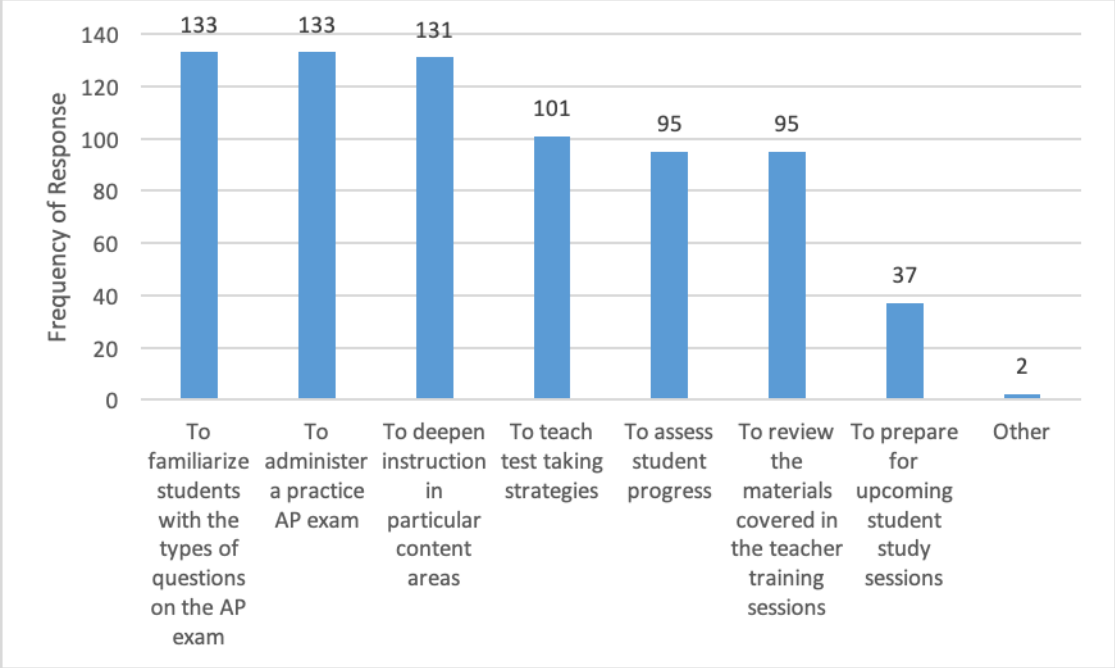



Figure 45. Ways in which survey participants incorporate CRP resources into instruction.

When asked how instructional resources had changed year-over-year, the most common response was that new resources were available (see Figure 46).

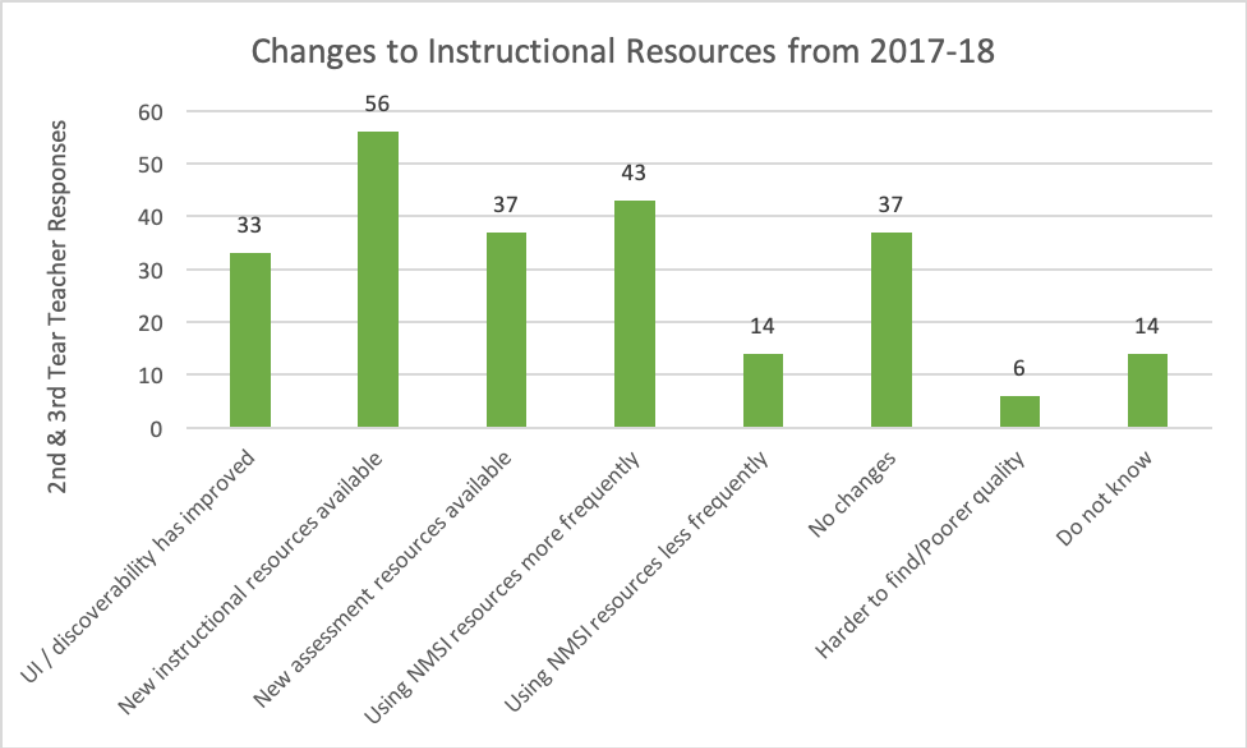



Figure 46. Changes to CRP instructional resources as noted by teachers in the 2nd and 3rd year of the CRP.

Third year teachers at Treatment Schools who responded to the survey in both 2018 and 2019 were more likely to say in their second year that they were using CRP resources more frequently than the previous year (see Table 78).

Table 78

*Third-Year Teacher CRP Resources Changes 2018 and 2019 Surveys (*n = *63)*

| **Changes to NMSI Resources Y-O-Y** | **2018 Survey** ***M (SD)*** | **2019 Survey** ***M (SD)*** | **2018–2019 diff** | **S.E.** | **t** | **df** | ***p*** |
|---|---|---|---|---|---|---|---|
| New/different instructional resources available | 0.5 (0.504) | 0.43 (0.50) | 0.071 | 0.077 | 0.652 | 59 | 0.517 |
| I am using NMSI / CRP resources more frequently | 0.45 (0.50) | 0.33 (0.48) | 0.117 | 0.06 | 2.206 | 59 | 0.031 |
| New/different assessment resources available | 0.42 (0.50) | 0.32 (0.47) | 0.1 | 0.083 | 1 | 59 | 0.321 |
| User interface / discoverability has improved | 0.25 (0.44) | 0.25 (0.44) | -0.004 | 0.073 | -0.228 | 59 | 0.821 |
| No changes | 0.13 (0.34) | 0.18 (0.38) | -0.042 | 0.065 | -0.772 | 59 | 0.443 |
| Do not know | 0.1 (0.30) | 0.05 (0.22) | 0.052 | 0.05 | 1 | 59 | 0.321 |
| I am using NMSI / CRP resources less frequently | 0.03 (0.18) | 0.13 (0.34) | -0.094 | 0.039 | -2.56 | 59 | 0.013 |
| Other | 0.03 (0.18) | 0.06 (0.25) | -0.03 | 0.041 | -0.814 | 59 | 0.419 |

Survey participants were asked in a multiple choice, multiple selection item what additional tools and materials they would like to see from NMSI. By far, the most common response was structured AP curricular units (see Figure 47).

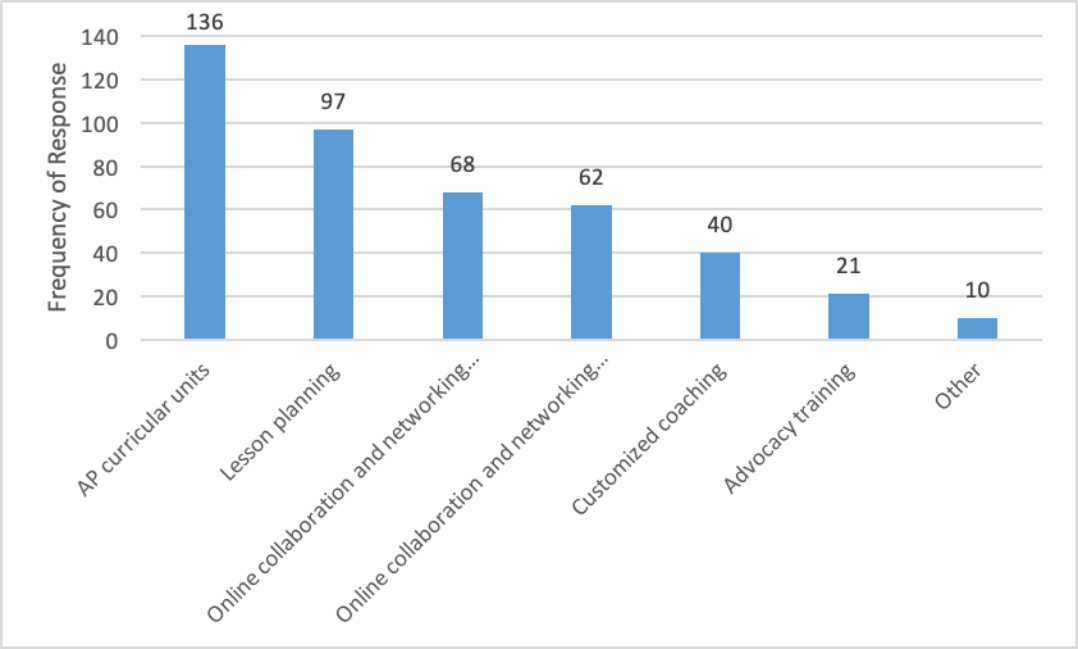



Figure 47. Additional tools and materials to which survey participants would like access through the CRP.

Respondents were given the option to provide further feedback, and comments ranged from wanting more guidance on scoring to wanting funds to purchase literature (see Appendix E).

### 7. 2018–2019 Teacher Survey: Goal Setting

Twenty-seven percent of survey participants ($n$ = 210) said that they met with NMSI representatives about exam score goals. When asked if their input was considered when establishing the goals, only 25% of 208 respondents said "yes." Asked to evaluate the CRP goals and the process for setting goals, survey participants agreed somewhat that the program established goals for equitable access to AP coursework (see Table 79).

Table 79

*Opinions about CRP Goals and the Goal Setting Process*

| Goal setting | *n* | *M* | *SD* |
|---|---|---|---|
| The program established goals for providing equitable access to AP coursework for all interested students | 207 | 2.90 | 0.82 |
| The program established measurable and attainable goals for class enrollment | 207 | 2.67 | 0.86 |
| The program established goals for recruitment of high-need and traditionally underrepresented students | 208 | 2.67 | 0.93 |
| The program established measurable and attainable goals for student exam performance | 207 | 2.65 | 0.91 |

*Note.* 1 = *strongly disagree*, 2 = *disagree somewhat*, 3 = *agree somewhat*, and 4 = *strongly agree*.

Figure 48 shows the results of comparing goal setting between the 2018 and 2019 survey. As one survey participant noted, "I will receive $1,000 if 21 of 11 students (yes that is the correct number) obtain a 3 or better. Might as well be $100,000,000."

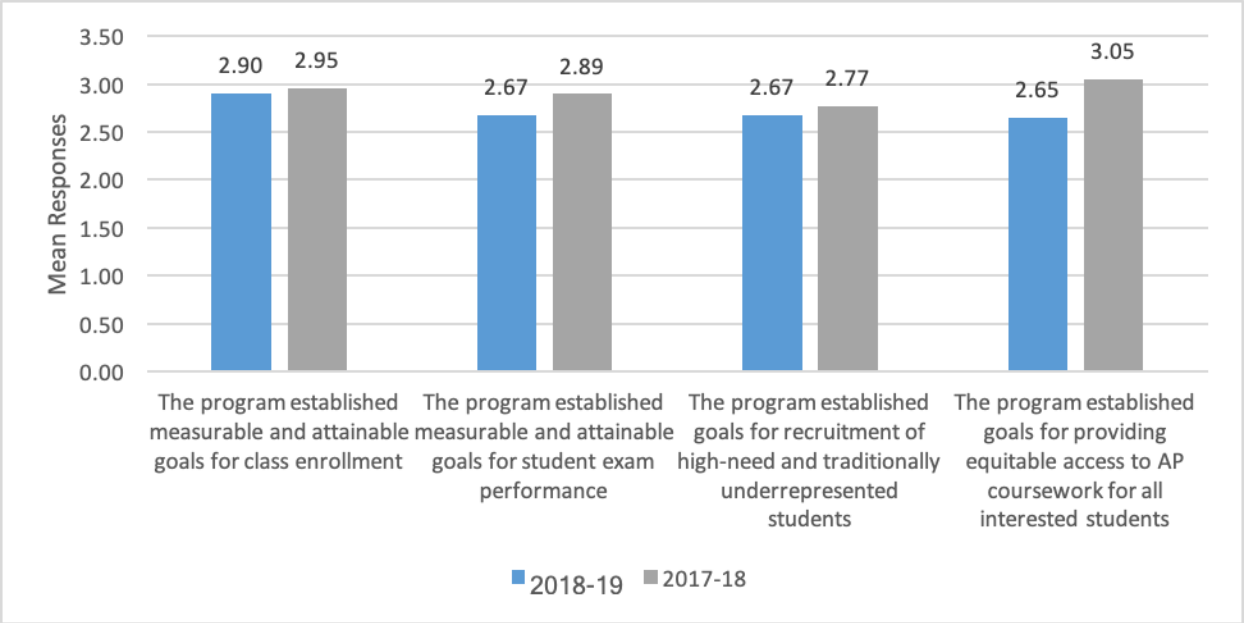



Figure 48. Comparison of 2018 survey results about goal setting with 2019 survey results.

While the level of agreement among third year teachers in Treatment Schools who completed surveys in 2018 and 2019 for all goal-related statements was lower in 2019, the only significant change concerned the attainability of qualifying score goals (see Table 80).

Table 80

*Third-Year Teacher CRP Goals 2018 and 2019 Surveys* (n = *63)*

| **Evaluation of Goals and Goal Setting** | **2018 Survey *M (SD)*** | **2019 Survey *M (SD)*** | **2018-2019 diff** | **S.E.** | **t** | **df** | ***p*** |
|---|---|---|---|---|---|---|---|
| The program established goals for providing equitable access to AP coursework for all interested students | 3.10 (0.71) | 2.95 (0.73) | 0.143 | 0.099 | 1.305 | 61 | 0.197 |
| The program established measurable and attainable goals for class enrollment | 3.03 (0.70) | 2.83 (0.81) | 0.207 | 0.113 | 1.818 | 62 | 0.074 |
| The program established measurable and attainable goals for student exam performance | 2.92 (0.79) | 2.68 (0.84) | 0.238 | 0.1 | 2.37 | 62 | 0.021 |
| The program established goals for recruitment of high-need and traditionally underrepresented students | 2.75 (0.79) | 2.78 (0.83) | 0.016 | 0.11 | 0.145 | 62 | 0.885 |

*Note.* 1 = *strongly disagree*, 2 = *disagree somewhat*, 3 = *agree somewhat*, and 4 = *strongly agree*.

## 8. 2018–2019 Teacher Survey: Overall Impact

Taken as a whole, the individual components of the CRP aim to push school cultures toward greater inclusion, higher expectations, and an emphasis on STEM education. Survey

participants felt that their schools' administration "promoted a culture of continuous improvement" and "valued STEM learning" at a higher rate than they felt the school set clear goals for either AP enrollment or exam performance (see Figure 49).

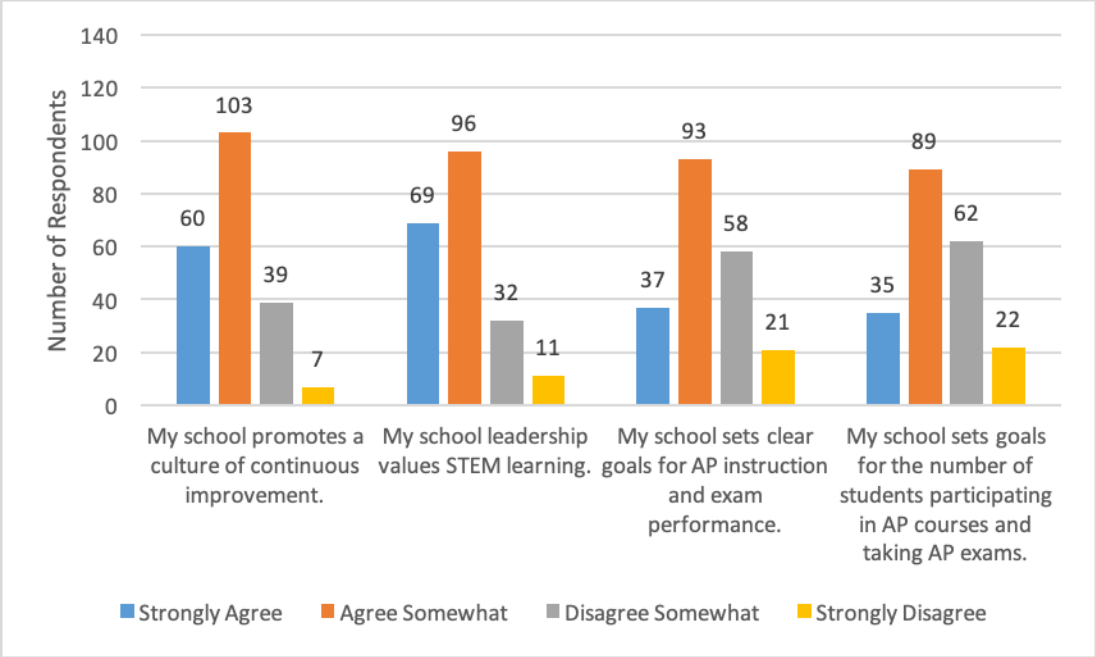



Figure 49. Administration leadership in school culture.

Perhaps because respondents felt strongly that their schools considered all students capable of achieving at high levels and encouraged all students to enroll in AP exams, they did not feel to a very strong extent that many students for whom AP was a good fit were being left behind (see Figure 50).

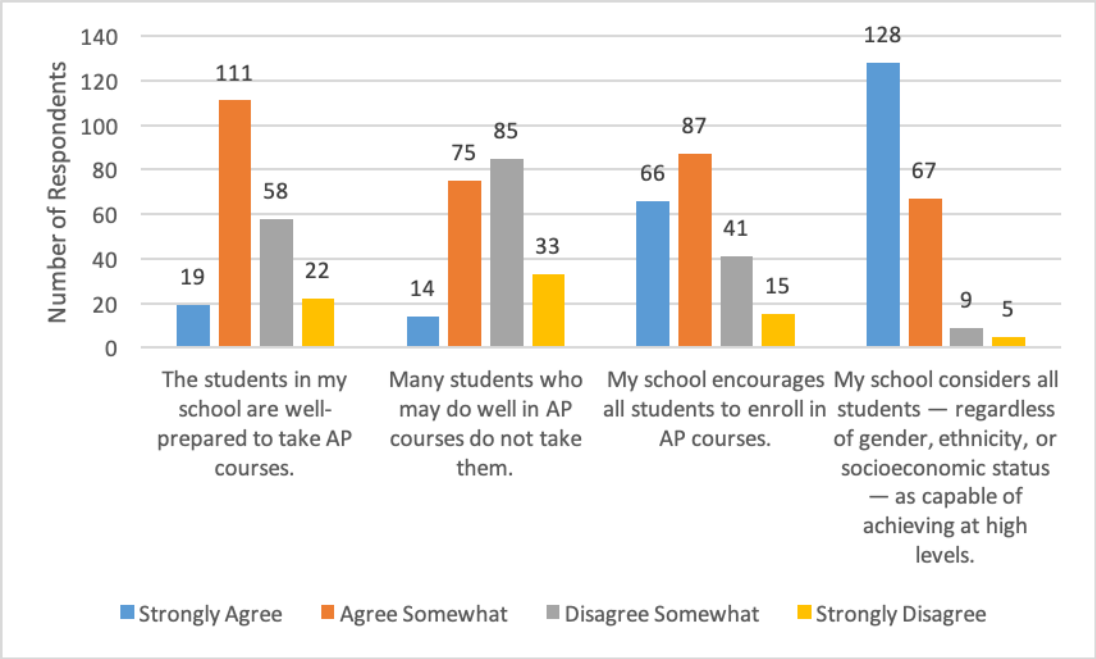



Figure 50. Opinions about student AP preparation.

In general respondents feel positively about students' preparation and students' confidence about their preparation for the year-end AP exam (see Figure 51).

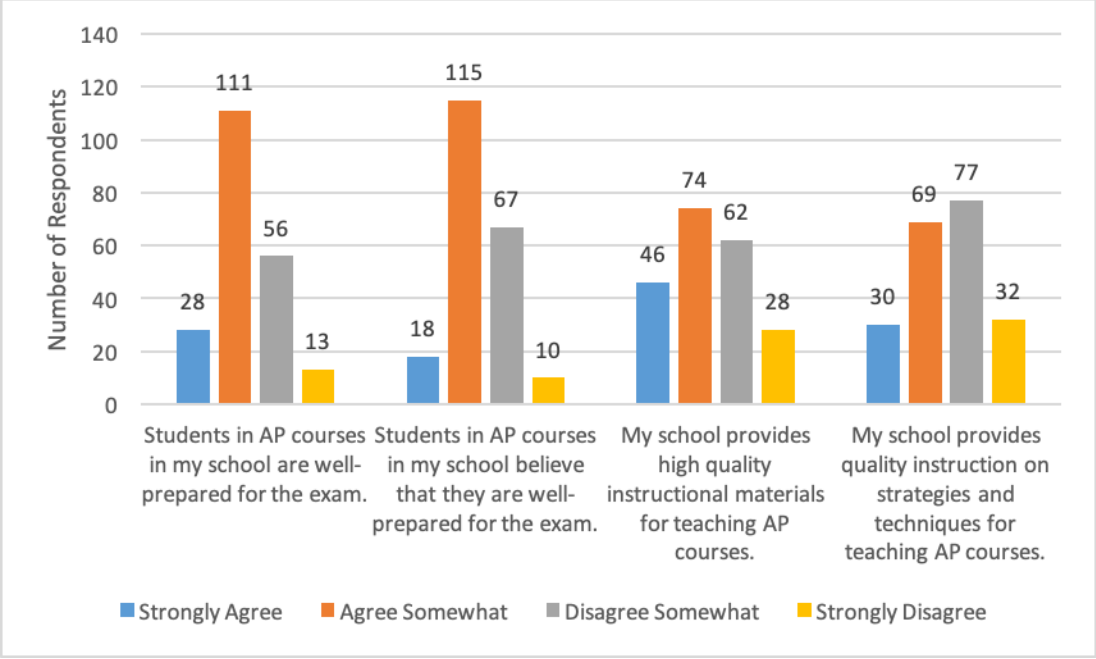



Figure 51. Impressions of the quality of AP instruction.

Respondents were asked to rate their agreement with a series of statements addressing the CRP's influence on their school's culture, and ratings were fairly consistent from 2018 to 2019 (see Figure 52).

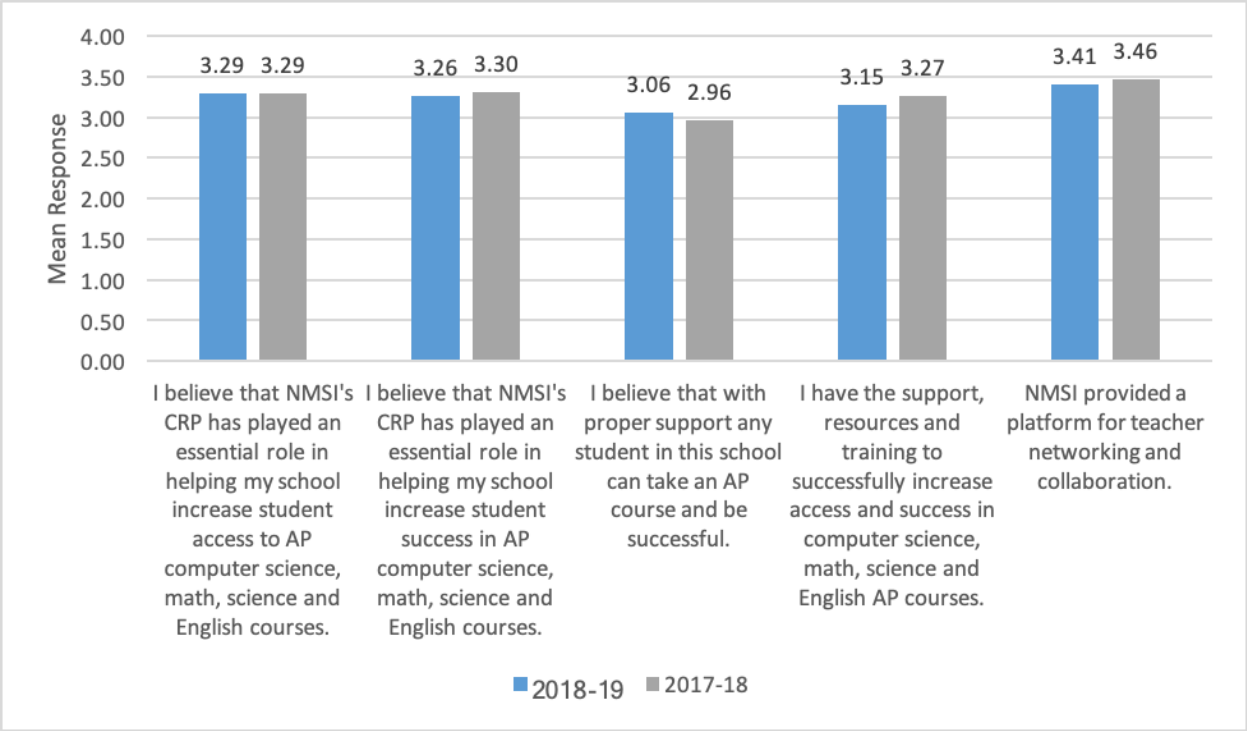



Figure 52. Comparison between 2018 and 2019 survey items about CRP influence on school culture.

When we compare ratings of second year teachers in Treatment Schools ($n$ = 32) to those from Delayed Treatment Schools ($n$ = 59), two significant and possibly related differences are revealed. To be clear, these are teachers with the same amount of time spent in the program but for one group the school is in its third year and for the other the school is in its second year. In one item, survey respondents from the Delayed Treatment Schools feel more strongly that their students put in the effort to excel in AP courses (see Table 81). In a separate item, the survey participants from Delayed Treatment Schools indicated that they believed more strongly that the School's AP students are well prepared for AP exams (see Table 82).

Table 81

*Second Year Teacher Comparison—Student Effort*

| School characteristics | Treatment *M (SD)* | Delayed Treatment *M (SD)* | Treatment-Delayed Treatment diff | S.E. | t | df | *p* |
|---|---|---|---|---|---|---|---|
| My students put in the effort it takes to learn in their AP course. | 2.60 (0.84) | 2.95 (0.71) | -0.355 | 0.174 | -2.041 | 55.102 | 0.046 |

*Note.* 1 = *strongly disagree*, 2 = *disagree somewhat*, 3 = *agree somewhat*, and 4 = *strongly agree*.

Table 82

*Second Year Teacher Comparison—Student Preparedness*

| School culture | Treatment *M (SD)* | Delayed Treatment *M (SD)* | Treatment-Delayed Treatment diff | S.E. | t | df | *p* |
|---|---|---|---|---|---|---|---|
| Students in AP courses in my school are well-prepared for the exam. | 2.50 (0.76) | 2.81 (0.74) | -0.341 | 0.166 | -2.061 | 62.246 | 0.043 |

*Note.* 1 = *strongly disagree*, 2 = *disagree somewhat*, 3 = *agree somewhat*, and 4 = *strongly agree*.

Comparing evaluations of school culture by second year teachers in Treatment Schools ($n$ = 32) to those of third year teachers in Treatment Schools ($n$ = 76), four significant differences emerge. In all four cases, the third year teachers rated the elements more highly than their second year counterparts (see Table 83).

Table 83

*Second and Third Year Teachers at Treatment Schools—Evaluation of School Culture*

| School culture | 2nd Year *M (SD)* | 3rd Year *M (SD)* | 2nd Yr-3rd Yr diff | S.E. | t | Df | *p* |
|---|---|---|---|---|---|---|---|
| My school provides quality instruction on strategies and techniques for teaching AP courses | 2.22 (0.83) | 2.76 (0.88) | -0.541 | 0.179 | -3.024 | 61.939 | 0.004 |
| Students in AP courses in my school believe that they are well-prepared for the exam | 2.44 (0.76) | 2.82 (0.67) | -0.378 | 0.155 | -2.448 | 52.173 | 0.018 |
| My school encourages all students to enroll in AP courses. | 2.69 (0.86) | 3.19 (0.87) | -0.499 | 0.182 | -2.746 | 59.008 | 0.008 |
| Students in AP courses in my school are well-prepared for the exam | 2.47 (0.76) | 2.79 (0.72) | -0.318 | 0.158 | -2.008 | 55.914 | 0.049 |

*Note.* 1 = *strongly disagree*, 2 = *disagree somewhat*, 3 = *agree somewhat*, and 4 = *strongly agree*.

The 61 Treatment School teachers who completed the survey each year felt more strongly after the third year that they had a good understanding of the concepts in their field when compared to the end of the first year of the program (see Table 84, Table 85, and Figure 53). On the other hand, these same survey participants felt less positively about student interest in learning after three years than they did after one year (see Table 84, Table 86, and Figure 54).

Table 84

*ANOVA—Teacher and Student Characteristics*

| Statement | 2016–2017 | | 2017–2018 | | 2018–2019 | | F (df) | pairs |
|---|---|---|---|---|---|---|---|---|
| | ***M*** | ***SD*** | ***M*** | ***SD*** | ***M*** | ***SD*** | | |
| I have a good understanding of the concepts I need to teach in my field. | 3.8 | 0.4 | 3.8 | 0.4 | 3.9 | 0.3 | 3.8 ( 2 ) | yr3>yr1 |
| I learn new ideas in my field quickly. | 3.7 | 0.5 | 3.7 | 0.5 | 3.7 | 0.5 | 0.3 ( 2 ) | |
| I have a strong sense of belonging to the community of educators. | 3.4 | 0.6 | 3.5 | 0.6 | 3.6 | 0.6 | 1.5 ( 2 ) | |
| My students are usually pretty interested in learning their AP course content. | 3.4 | 0.6 | 3.3 | 0.6 | 3.2 | 0.6 | 4.9 ( 2 ) | yr3<yr1 |
| My students put in the effort it takes to learn in their AP course. | 3.1 | 0.7 | 3.0 | 0.7 | 2.9 | 0.7 | 2.4 ( 2 ) | |

*Note.* 1 = *strongly disagree*, 2 = *disagree somewhat*, 3 = *agree somewhat*, and 4 = *strongly agree*.

Table 85

*Pairwise Comparison—Good Understanding of the Concepts I Need to Teach in My Field*

| contrast | estimate | std.error | df | t | p |
|---|---|---|---|---|---|
| yr2-yr1 | 0.08 | 0.05 | 120 | 1.52 | 0.13 |
| yr3-yr1 | 0.15 | 0.05 | 120 | 2.73 | 0.01 |
| yr3-yr2 | 0.07 | 0.05 | 120 | 1.21 | 0.23 |

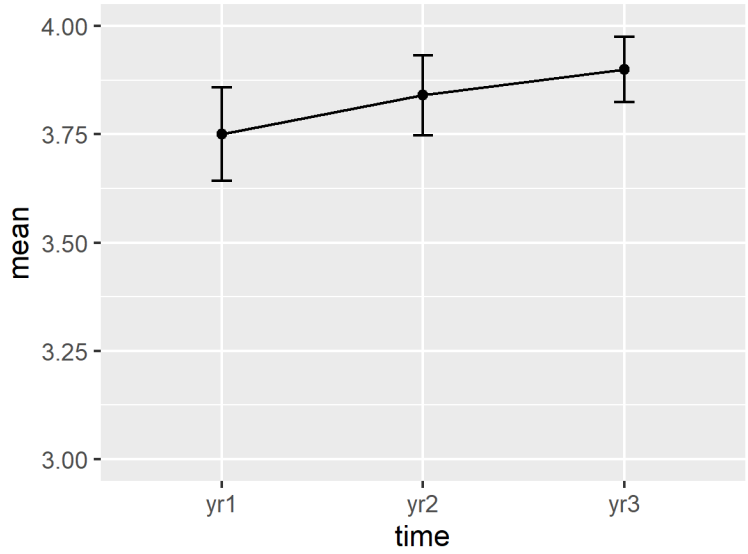



Figure 53. Change across three years—good understanding of the concepts I need to teach in my field.

Table 86

*Pairwise Comparison—Student Interest in Learning AP Content*

| contrast | estimate | std.error | df | t | p |
|---|---|---|---|---|---|
| yr2-yr1 | -0.10 | 0.08 | 120 | -1.25 | 0.22 |
| yr3-yr1 | -0.25 | 0.08 | 120 | -3.12 | 0.00 |
| yr3-yr2 | -0.15 | 0.08 | 120 | -1.87 | 0.06 |

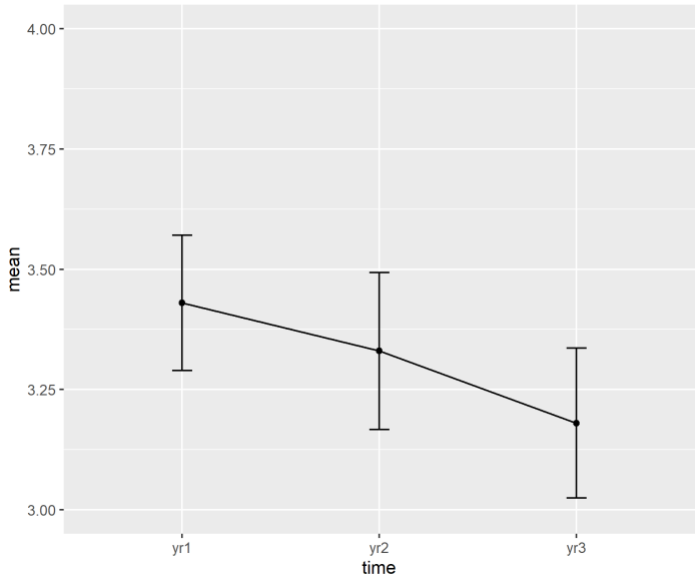



Figure 54. Change across three years—student interest in learning AP content.

One of the goals of the CRP is to increase AP enrollment in participating schools and a key tenet is that all AP courses should be open to all students. Figure 55 shows that 75% of survey participants felt that the CRP is an effective means of increasing AP enrollment.

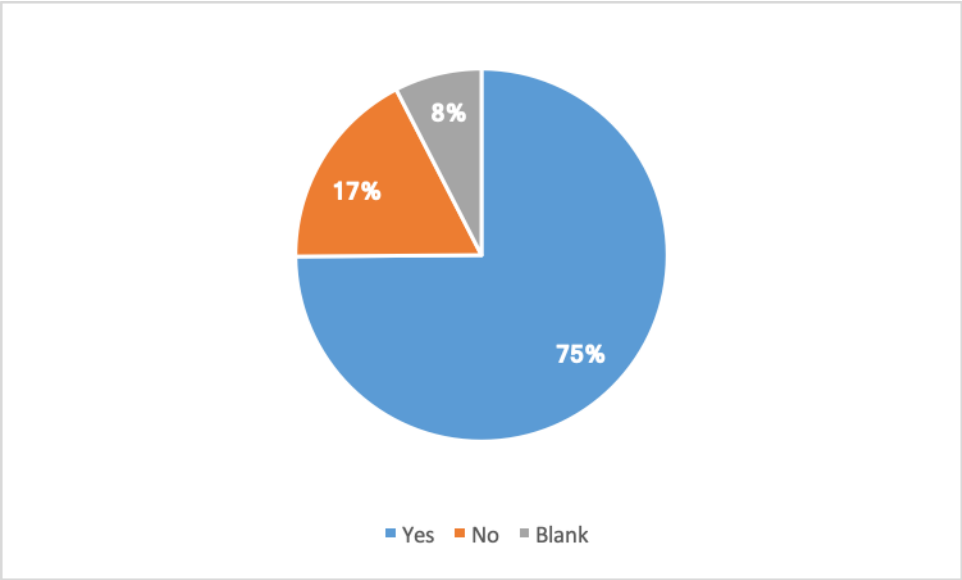



Figure 55. Impact of CRP on AP enrollment.

Survey participants took the opportunity to provide feedback on the program's impact on AP enrollment. Comments ranged from specific logistics in the respondent's school to musings about the philosophy of open enrollment (see Appendix F).

When asked if open enrollment had a positive effect on the AP program at their school, 74% of survey participants said yes (see Figure 56).

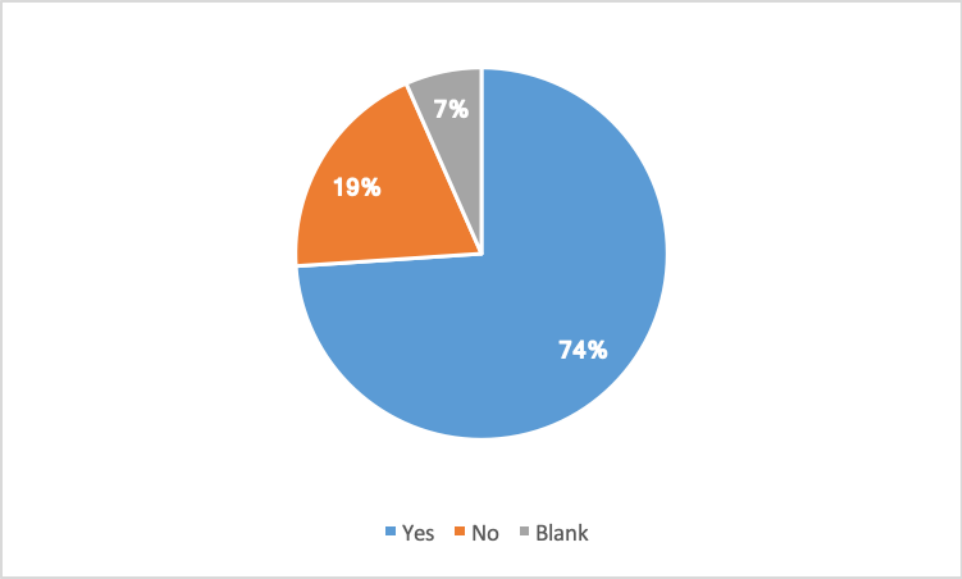



Figure 56. Percentage of teachers who felt that open enrollment had a positive effect on AP program.

Survey participants were asked if the CRP contributed to improving each of the seven stated goals of the CRP. As the data in Table 87 show, respondents felt that the CRP did contribute to improvements in all seven areas.

Table 87

*CRP Contributions to Improvements in Key Areas*

| College Readiness Program goal | *n* | *M* | *SD* |
|---|---|---|---|
| Teachers' instructional skills, techniques and strategies | 201 | 2.55 | 0.59 |
| Teachers' content knowledge | 200 | 2.52 | 0.60 |
| Students' content knowledge | 199 | 2.39 | 0.65 |
| Students' experience with STEM AP courses | 200 | 2.32 | 0.67 |
| Recruitment of high-need and traditionally underrepresented students into AP courses | 201 | 2.21 | 0.72 |
| School culture of continuous improvement | 200 | 2.14 | 0.69 |
| School leadership valuing STEM learning | 200 | 2.14 | 0.73 |

*Note.* 1 = *no improvement*, 2 = *slight improvement*, 3 = *major improvement.*

Among the 61 teachers who completed the survey all three years, we see a steady increase in their belief that the CRP is an effective way to increase student enrollment in AP courses (see Table 88).

Table 88

*"Do you feel that the CRP is an effective way to increase student enrollment in AP courses?"*

| | 2016–2017 | | 2017–2018 | | 2018–2019 | | sig |
|---|---|---|---|---|---|---|---|
| **Response** | **Freq** | **%** | **Freq** | **%** | **Freq** | **%** | **Chi (df)** |
| Yes=1 | 47 | 79.7 | 50 | 82.0 | 51 | 86.4 | 1.2 ( 2 ) |

Specifically, the 63 third-year teachers who completed a survey in 2018 and in 2019 felt more strongly that the CRP improved the recruitment of underrepresented students into AP courses in 2019 than they did in 2018 (see Table 89).

Table 89

*Third-Year Teacher Recruitment of High Need Students 2018 and 2019 Surveys (*n = *63)*

| **Improvement in seven target areas** | **2018 Survey** ***M (SD)*** | **2019 Survey** ***M (SD)*** | **2018-2019 diff** | **S.E.** | **t** | **df** | ***p*** |
|---|---|---|---|---|---|---|---|
| Recruitment of high-need and traditionally underrepresented students into AP courses | 2.079 (0.725) | 2.254 (0.718) | -0.175 | 0.086 | -2.024 | 62 | 0.047 |

*Note.* 1 = *no improvement*, 2 = *slight improvement*, 3 = *major improvement.*

## 9. 2018–2019 Teacher Survey: Effective Components

Survey participants were asked in three distinct items to select from a list of CRP elements only one *most* effective, *second most* effective, and *least* effective CRP component. As Figure 57 illustrates, 75% of teachers responding to the items rank teacher training as either the most or second most effective CRP component.

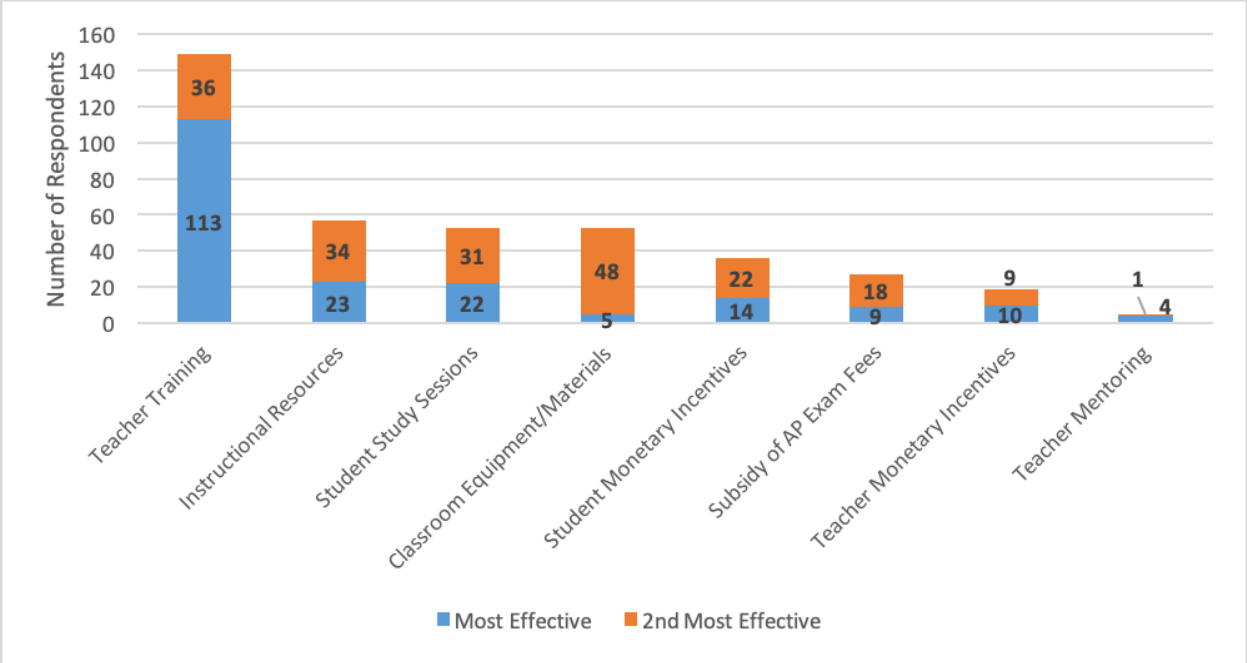



Figure 57. Survey participant choices for most and second most effective CRP components.

The mentoring program was viewed as the least effective component of the CRP by 33% of the respondents to the item (see Figure 58).

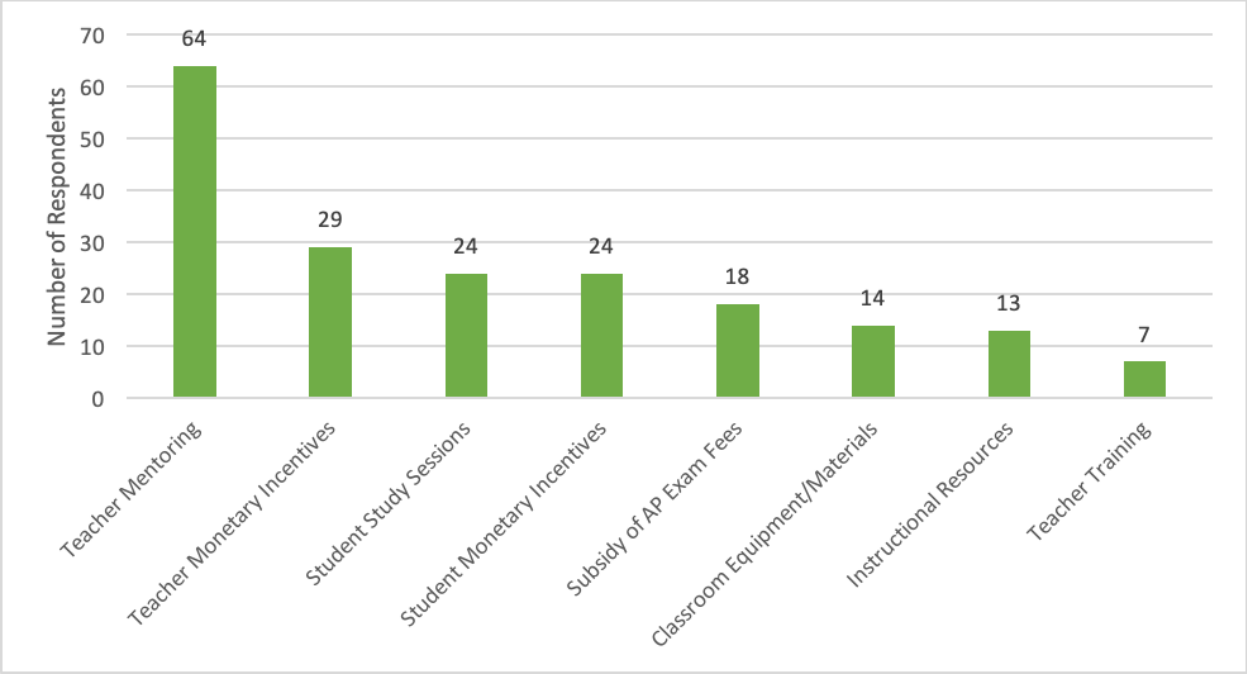



Figure 58. Survey participant choices for least effective CRP components.

### b. 2018–2019 Partner School Director Survey Response Summary

In April 2019, electronic surveys were disseminated to the Partner School Directors (PSDs) in all 48 schools participating in the study. The CRP seeks to expand access to AP courses for underserved populations, and the PSDs are the administrators in each school taking the lead on the AP expansion. The surveys are intended to gather information beyond the participation and test score data available elsewhere. The goal of the surveys (and select follow-up interviews) is to gain a deeper understanding of the history of AP instruction at the school, the successes and challenges in program implementation, and the perceptions of the efficacy of the program.

In 2019, surveys were completed by 28 PSDs representing 28 schools from ten states and 11 metropolitan areas—at least one school from each of the regions covered by the study. Completed surveys were received from 13 Treatment Schools and 15 Delayed Treatment Schools. PSDs who completed the surveys had been in the role for one ($n$ = 7), two ($n$ = 13), or three years ($n$ = 8).

PSDs were asked if they agreed—on a 4-point scale from *strongly disagree* (1) to *strongly agree* (4)—that the CRP matched their expectations for the 2018–2019 school year. As Figure 59 shows, 82% of survey participants agree to some degree that the program met expectations.

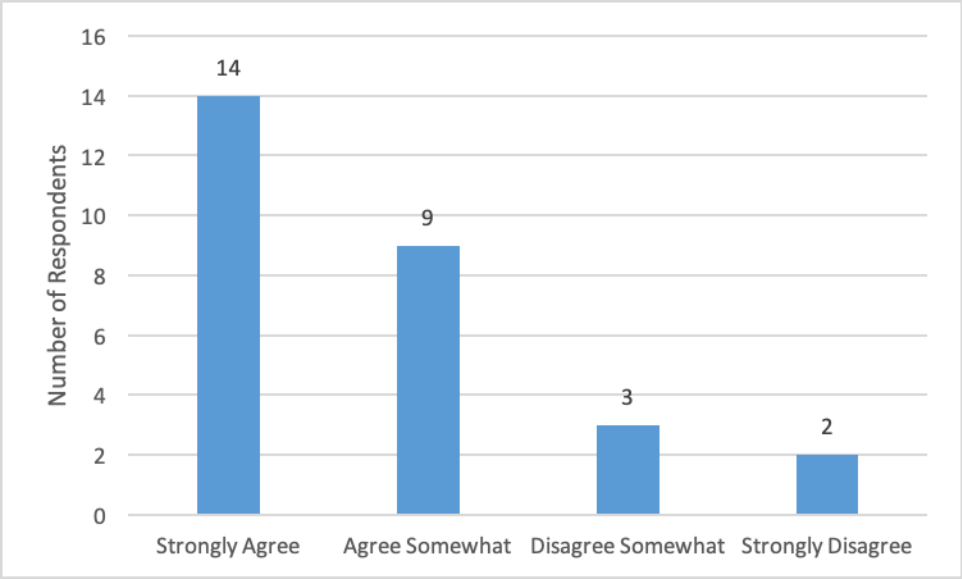



Figure 59. Degree to which CRP matched expectations.

PSDs were asked to elaborate on the ways in which the CRP did or did not match their expectations, and their responses comprise Appendix G.

Survey participants were asked the extent to which they agree with a series of statements about the philosophies behind the CRP and the implementation of the CRP in their school. Most teachers agree or strongly agree with each of the statements (see Table 90). PSDs provided additional feedback on the CRP's impact in these areas, and the comments comprise Appendix H.

Table 90

*NMSI's Role and Philosophy*

| NMSI impact and philosophy | *n* | *M* | *SD* |
|---|---|---|---|
| I believe that with proper support any student in this school can take an AP course and be successful. | 27 | 3.52 | 0.63 |
| I have the support, resources and training to successfully increase access and success in computer science, math, science and English AP courses. | 28 | 3.46 | 0.63 |
| NMSI provided a platform for networking and collaboration. | 27 | 3.44 | 0.74 |
| I believe that NMSI's CRP has played an essential role in helping my school increase student success in AP computer science, math, science and English courses. | 28 | 3.43 | 0.82 |
| I believe that NMSI's CRP has played an essential role in helping my school increase student access to AP computer science, math, science and English courses. | 28 | 3.39 | 0.82 |

*Note.* 1 = *strongly disagree*, 2 = *disagree somewhat*, 3 = *agree somewhat*, and 4 = *strongly agree*.

Only three of 28 PSDs surveyed said that the CRP was not an effective way to increase enrollment. One of those three further clarified that adding courses under the auspices of the program increased the number of students taking AP courses, but the CRP did not increase enrollment in existing AP courses. Only one respondent said that open enrollment did not have a positive impact on the AP program at their school.

Survey participants were asked to select from a list of eight program components, the most and "second most" effective elements of the CRP. Consistent with the teacher survey results, many respondents (21) felt that teacher training was one of the more effective components of the program (see Figure 60).

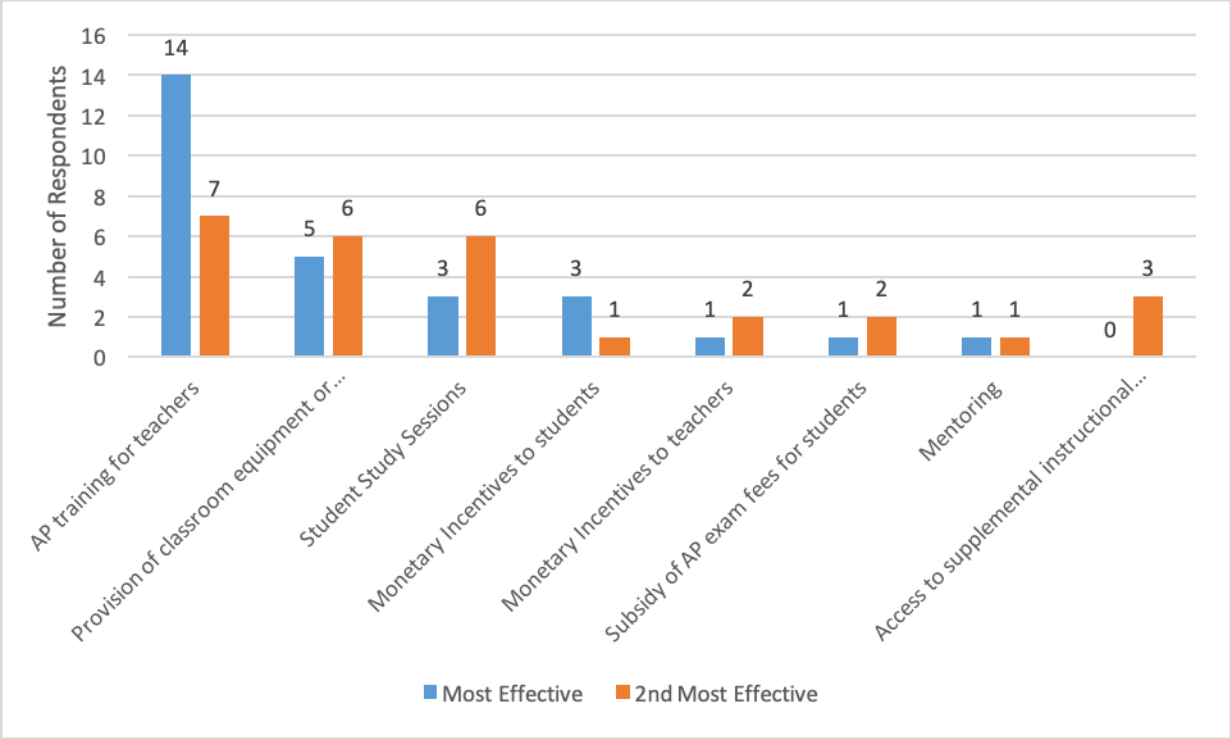



Figure 60. Effective components of the CRP.

Respondents were also asked which CRP component was the least effective. As it did in the teacher survey, mentoring topped the list of least effective components. However, it is unclear if mentoring was ineffective in implementation—meaning that teachers were not successfully matched to mentors—or ineffectual for those who were successfully matched to mentors (see Figure 61).

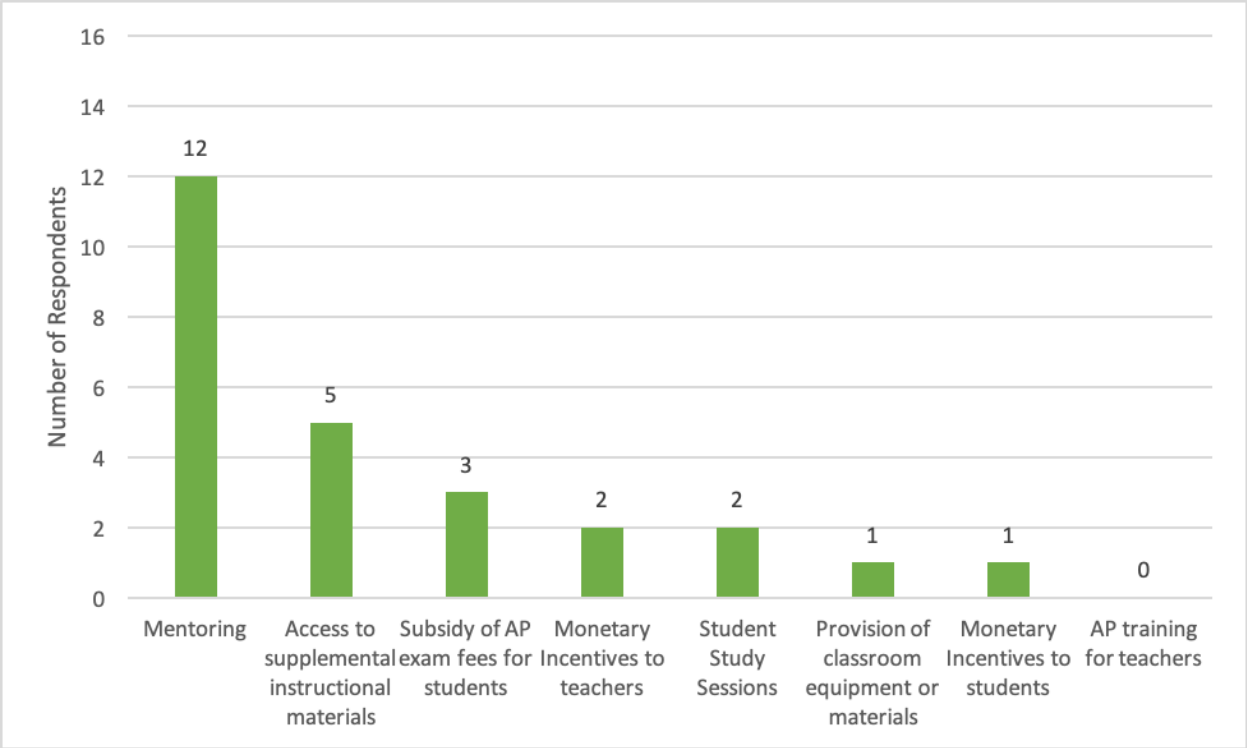



Figure 61. Least effective component of the CRP.

PSDs were asked if the CRP contributed to improving seven characteristics of the school, each tied to a stated goal of the CRP. The improvement to each characteristic was rated on a 3-point scale: *no improvement* (1), *slight improvement* (2), and *major improvement* (3). More survey respondents felt that the CRP improved the students' experience with STEM AP courses and the instructional skills of teachers (see Table 91).

Table 91

*CRP Contributions to School Improvements*

| School activities and characteristics | *n* | *M* | *SD* |
|---|---|---|---|
| Students' experience with STEM AP courses | 28 | 2.68 | 0.47 |
| Teachers' instructional skills, techniques and strategies | 28 | 2.64 | 0.48 |
| Teachers' content knowledge | 28 | 2.61 | 0.49 |
| Students' content knowledge | 28 | 2.57 | 0.49 |
| Recruitment of high-need and traditionally underrepresented students into AP courses | 28 | 2.46 | 0.63 |
| School culture of continuous improvement | 28 | 2.46 | 0.73 |
| School leadership valuing STEM learning | 28 | 2.43 | 0.68 |

*Note.* 1 = *no improvement*, 2 = *slight improvement*, 3 = *major improvement.*

Survey participants' comments about factors that contributed to the success of the CRP in their school, or impeded implementation in their school comprise Appendix I.

### c. 2018–2019 Site Coordinator Survey Response Summary

The responsibilities of the Site Coordinator vary from school to school, but in all cases the Site Coordinator is the point person for student study session organization and registration. The Site Coordinator survey instrument reflects this consistent role. While many items are similar to the teacher and Partner School Director survey items, the Site Coordinator item may deviate slightly. As an example, when asked about Program Manager support, only Site Coordinators were given a "don't know" option.

Surveys were distributed to all Site Coordinators in each of 48 schools in the study, including schools with more than one Site Coordinator. Completed surveys were returned by 36 Site Coordinators representing 16 Delayed Treatment Schools and 20 Treatment Schools. Twelve SCs were in their third year, 19 in their second, and 5 in their first year coordinating the CRP.

Twenty of the 36 Site Coordinators said that their schools provided transportation to student study sessions. Only five respondents provided any further detail about the transportation offered to students: four schools provided bus passes to students and the fifth arranged for school buses. Other Site Coordinator suggestions for improving student attendance at sessioins comprise Appendix J.

Participants were asked to respond to six qualitative statements about the student study sessions. Across the board, Site Coordinators hold the sessions in high regard (see Table 92).

Table 92

*Qualitative Assessment of SSS*

| Student study session observation | *n* | *M* | *SD* |
|---|---|---|---|
| The study sessions were led by AP experts who taught NMSI-created lessons | 36 | 3.69 | 0.52 |
| The study sessions improved the students' content knowledge | 36 | 3.56 | 0.50 |
| Students were active participants (e.g., answering and asking questions, focused on tasks assigned, etc.) | 36 | 3.47 | 0.60 |
| The study sessions helped to increase student confidence | 36 | 3.47 | 0.64 |
| The study sessions highlighted the instructional needs for teachers to continue addressing in class | 36 | 3.44 | 0.72 |
| The study sessions were conveniently scheduled to accommodate student schedules | 36 | 3.25 | 0.68 |

*Note.* 1 = *strongly disagree*, 2 = *disagree somewhat*, 3 = *agree somewhat*, and 4 = *strongly agree*.

Site coordinators were asked if they thought that the study sessions were useful both for students and teachers. All but one respondent felt that the sessions were useful or extremely useful for students. They were less certain about the benefit to teachers (see Figure 62).

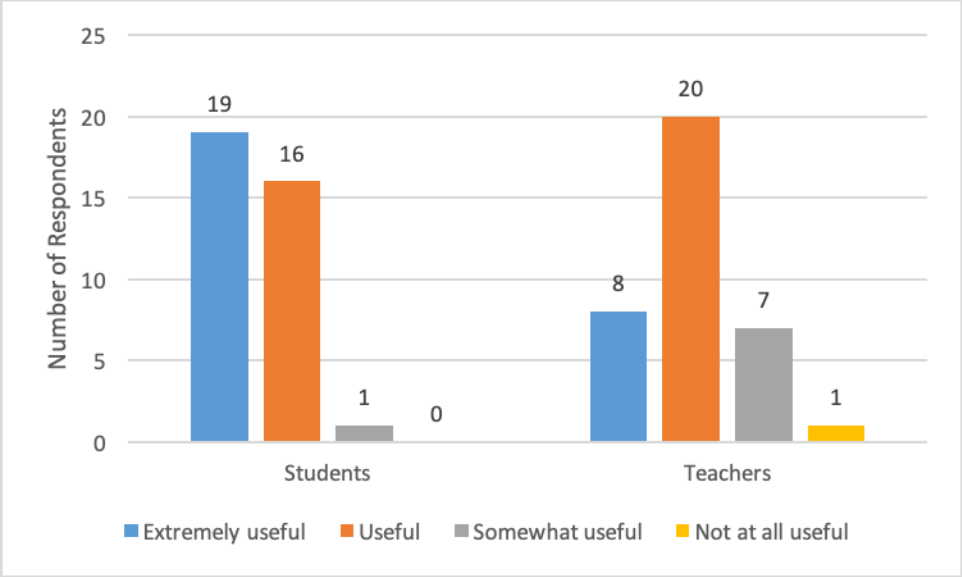



Figure 62. Rating the efficacy of the student study sessions.

Site coordinators were asked about the CRP's impact on AP enrollment. Thirty-two survey participants believe that the CRP is an effective way to increase AP enrollment, and 33 believe that open AP enrollment had a positive impact on the AP program at their school.

The CRP targets improvements in seven key areas, and the Site Coordinators were asked to evaluate whether the program made no improvement (1), a slight improvement (2), or a major improvement in each of the areas. Respondents felt that the CRP had the most impact on students' experience with STEM AP courses and made the least improvement in the school's recruitment of high-need students into AP courses (see Table 93).

Table 93

*Rating Impact on Areas Targeted by CRP*

| Did CRP improve the following? | *n* | *M* | *SD* |
|---|---|---|---|
| Students' experience with STEM AP courses | 33 | 2.52 | 0.50 |
| Students' content knowledge | 34 | 2.47 | 0.50 |
| Teachers' content knowledge | 34 | 2.47 | 0.55 |
| School leadership valuing STEM learning | 35 | 2.43 | 0.60 |
| Teachers' instructional skills, techniques and strategies | 33 | 2.39 | 0.65 |
| School culture of continuous improvement | 34 | 2.38 | 0.73 |
| Recruitment of high-need and traditionally underrepresented students into AP courses | 35 | 2.29 | 0.70 |

*Note.* 1 = *no improvement*, 2 = *slight improvement*, 3 = *major improvement.*

Site coordinators were given the opportunity to provide insights into specific challenges they faced implementing the CRP or factors that helped with implementing the CRP. Respondents' comments comprise Appendix K.

### d. 2018–2019 Student Survey Response Summary

Toward the end of the 2018–2019 school year, 6,380 paper student surveys were distributed to 160 teachers participating in NMSI's College Readiness Program. Ninety-three teachers returned a total of 2,710 completed surveys, for a return rate of 42%. The surveys represent students in 38 of the 48 schools in the study from all ten states in which study schools are located. More than half of the responses came from Georgia, North Dakota, and Texas (see Table 94).

Table 94

*Participant Count by State*

| State | Number of surveys | Number of schools | Schools w/ surveys |
|---|---|---|---|
| California | 121 | 2 | 2 |
| Georgia | 458 | 4 | 4 |
| Illinois | 333 | 5 | 4 |
| Louisiana | 112 | 2 | 2 |
| Michigan | 211 | 3 | 2 |
| Missouri | 148 | 6 | 4 |
| North Dakota | 502 | 5 | 5 |
| Ohio | 79 | 8 | 4 |
| Pennsylvania | 320 | 5 | 5 |
| Texas | 426 | 8 | 6 |
| TOTAL | 2710 | 48 | 38 |

Responses from Treatment Schools comprised 65% of total responses (see Table 95).

Table 95

*Distribution of Survey Responses by School Group*

| School group | Survey count | % of surveys | School count | % of schools |
|---|---|---|---|---|
| Treatment | 1768 | 65% | 22 | 58% |
| Delayed Treatment | 942 | 35% | 16 | 42% |

### 1. 2018–2019 Student Survey: AP Courses

In the 2019 Survey, 48% of students said that they were taking only one AP STEM or ELA course, which was an increase from the 2018 Survey (37%). See Figure 63.

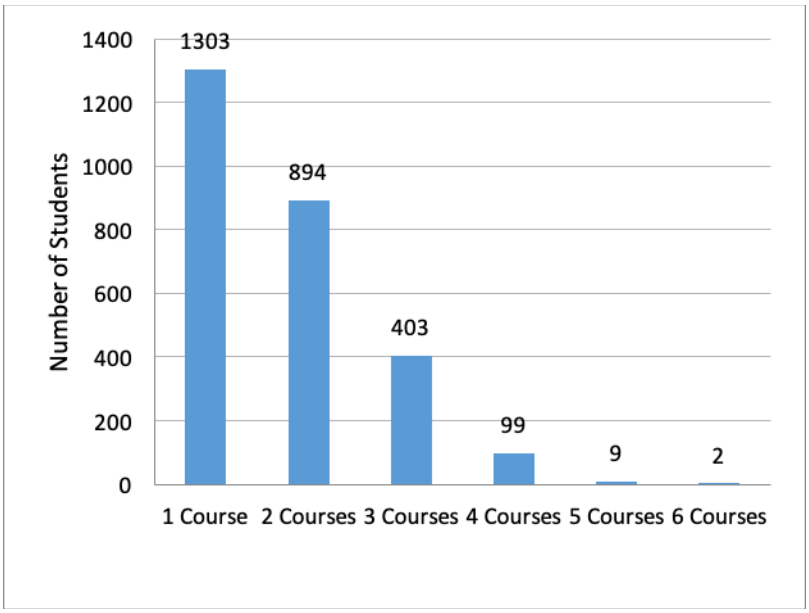



Figure 63. Respondent STEM/ELA AP courseload.

A greater proportion of Delayed Treatment participants are taking only one course when compared to Treatment School respondents, however a higher percentage are also taking three courses (see Figure 64).

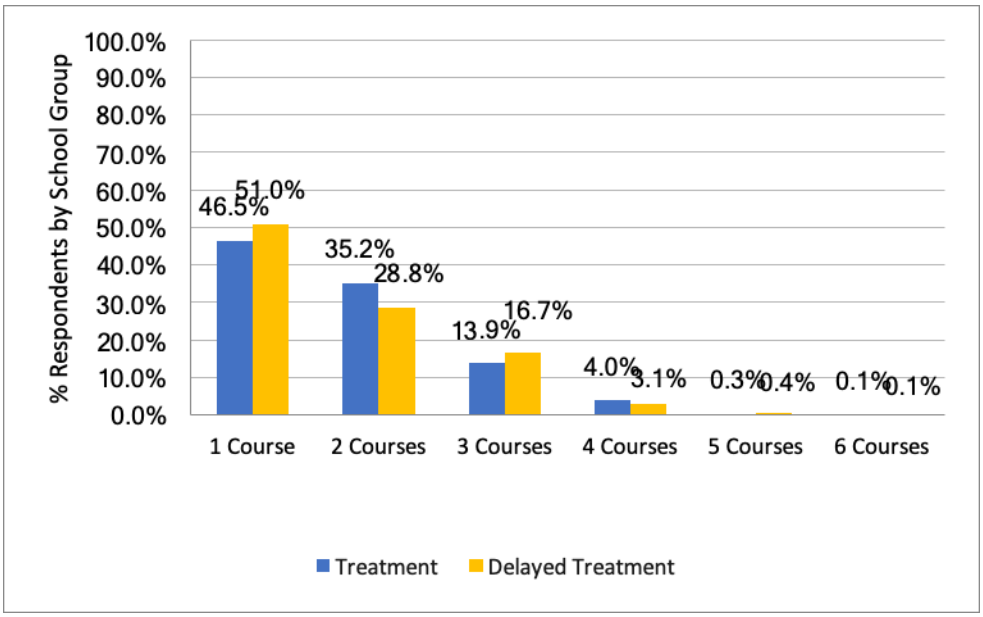



Figure 64. Number of courses taken by school group.

Each course subject is represented in the participant pool (see Figure 65).

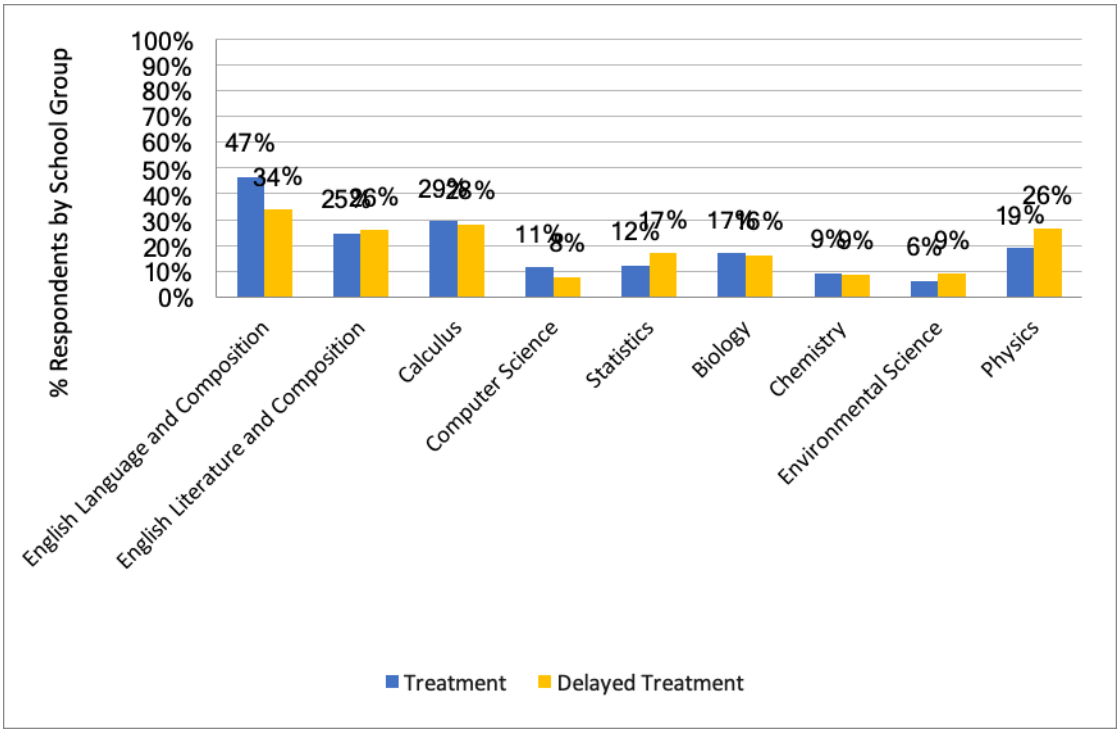



Figure 65. Percentage of participants enrolled in AP subjects by school group.

Of the courses included in the study, 40% of the Delayed Treatment participants are enrolled exclusively in STEM AP courses, as opposed to 29% of Treatment School respondents (see Figure 66).

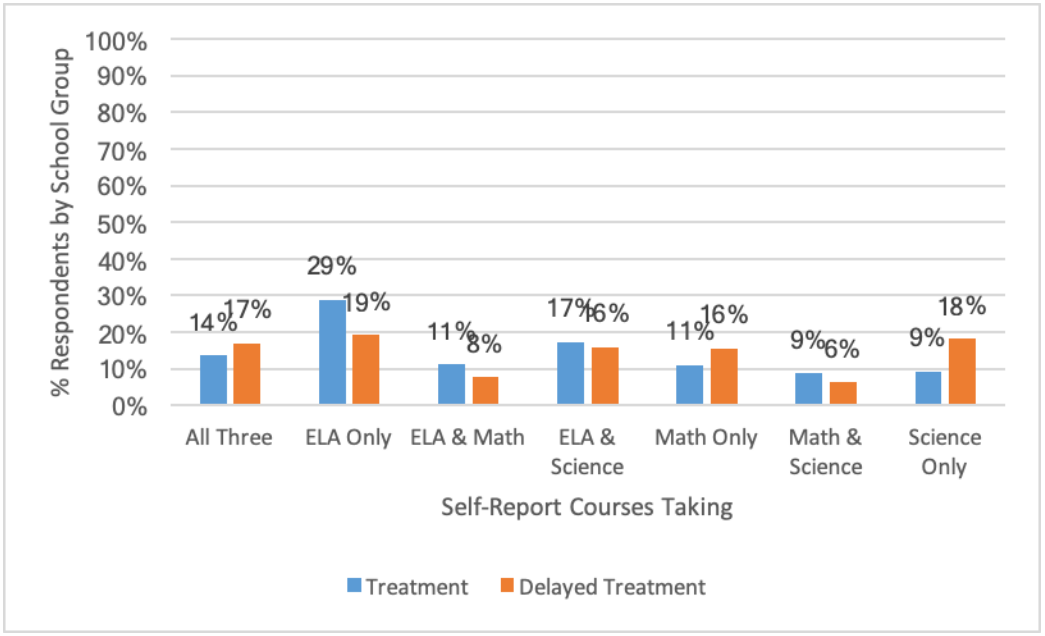



Figure 66. Breakdown of course enrollment by school group.

### 2. 2018–2019 Student Survey: Knowledge of AP and the CRP

Sixty-four percent of participants said they learned about AP courses from more than one source, which may be an indication of a school's commitment to increasing AP enrollment. Figure 67 illustrates survey participants' sources of information about AP courses.

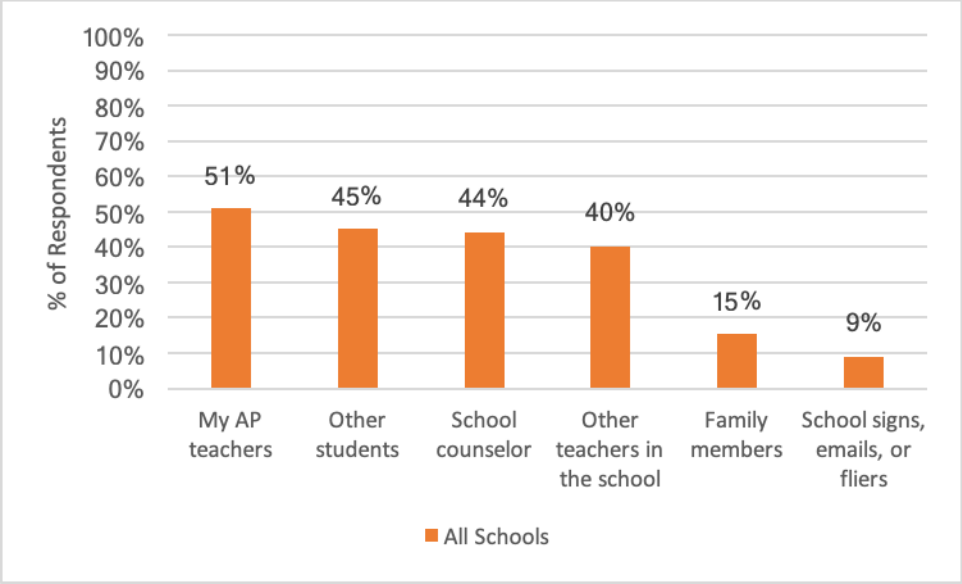



Figure 67. Ways in which participants learned about AP courses.

### 3. 2018–2019 Student Survey: Future Educational Plans

The vast majority of survey respondents (87%) plan to attend at least a four-year college or university (see Figure 68).

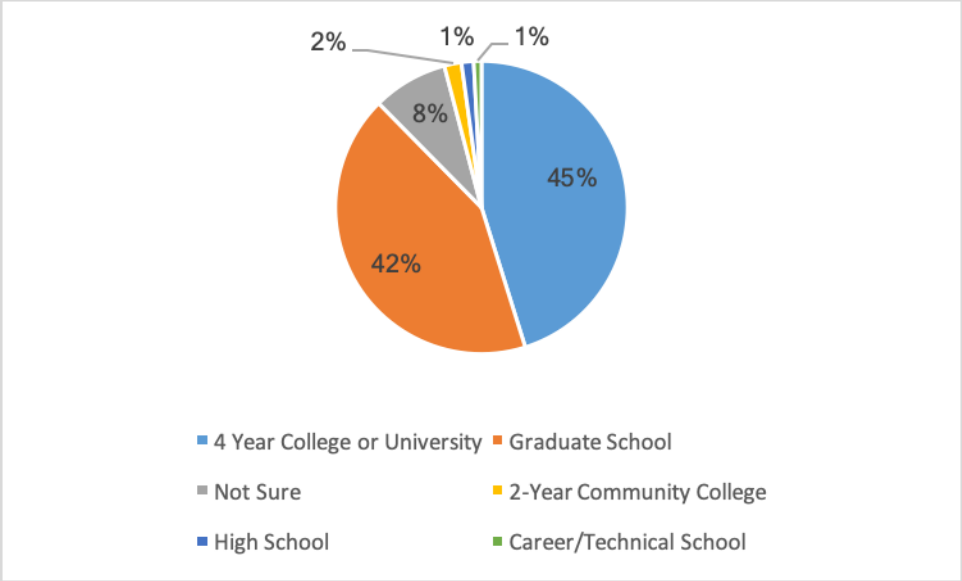



Figure 68. Level of education participant expects to achieve.

### 4. 2018–2019 Student Survey: Student Study Sessions

Survey participants were asked several questions about the student study sessions for the course in which they were completing the survey. Nearly all students learned about the Student Study Sessions from their AP teachers (see Table 96).

Table 96

*Ways in Which Participants Learned About Student Study Sessions*

| How student learned about student study sessions | *n* | % |
|---|---|---|
| My AP teachers | 2,613 | 97% |
| Other students | 457 | 17% |
| School counselor | 278 | 10% |
| Other teachers in the school | 183 | 7% |
| School signs, emails, or fliers | 207 | 8% |
| Other | 29 | 1% |

At the beginning of the school year, three sessions are scheduled for each course. However, due to scheduling conflicts or weather issues a session may be canceled and not rescheduled. Survey respondents indicated that not all of their courses had three sessions (see Table 97).

Table 97

*Number of Student Study Sessions Offered for Course*

| Number of sessions | N | % |
|---|---|---|
| 0 | 72 | 3% |
| 1 | 34 | 1% |
| 2 | 255 | 10% |
| 3 | 2,304 | 86% |

Survey participants from Treatment Schools attended student study sessions at a higher rate than respondents from Delayed Treatment Schools (see Figure 69). Twenty-six percent of respondents from Treatment Schools and 21% of Delayed Treatment School survey participants attended 100% of the available sessions.

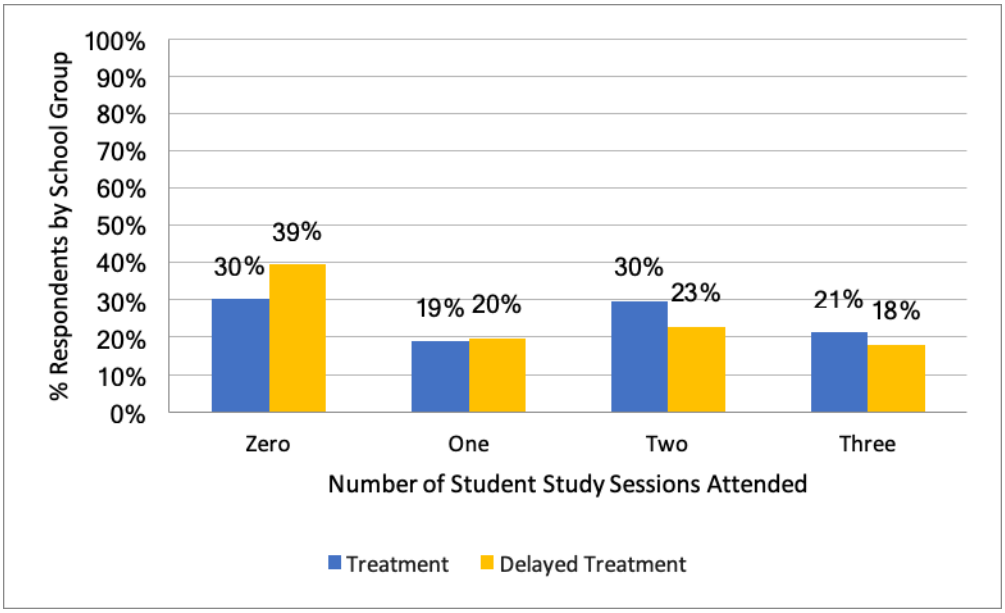



Figure 69. Number of sessions attended by school group.

Treatment Schools provide transportation and opportunities to make up sessions at a higher rate than Delayed Treatment Schools (see Figure 70).

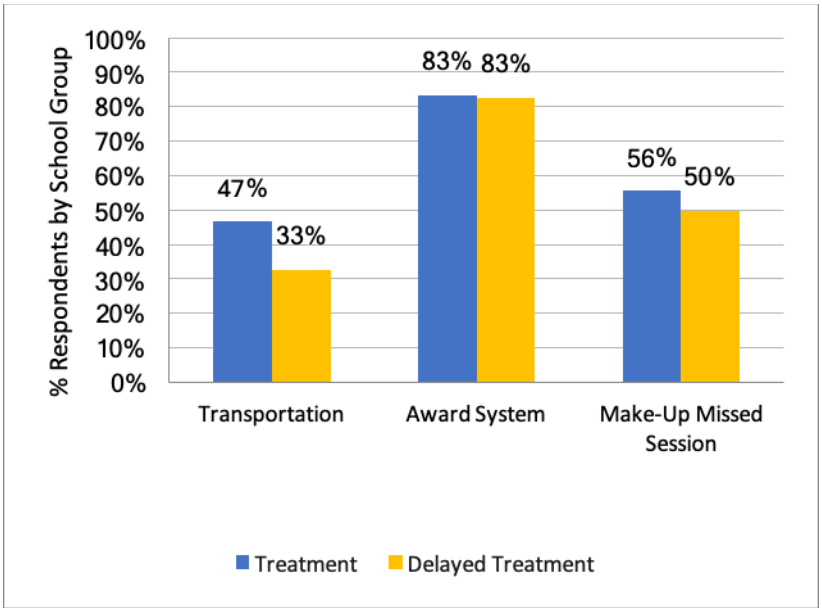



Figure 70. School actions in support of student study sessions.

## 5. 2018–2019 Student Survey: Student Opinions about Study Sessions

Survey participants felt that the student study sessions improved their content knowledge and increased their confidence in their own ability to complete the course successfully (see Table 98).

Table 98

*Evaluation of Student Study Sessions*

| Description of student study session | *n* | *M* | *SD* |
|---|---|---|---|
| Student Study Sessions improved my understanding of the course content | 1,803 | 3.22 | 0.65 |
| Student Study Sessions increased my confidence in my ability to successfully complete the AP course | 1,803 | 3.11 | 0.66 |
| Student Study Sessions improved my test taking strategies | 1,800 | 3.05 | 0.76 |
| Student Study Sessions increased my confidence in my ability to get a score of 3 or more on the AP exam | 1,805 | 2.99 | 0.72 |
| Student Study Sessions were conveniently scheduled | 1,795 | 2.92 | 0.84 |
| Student Study Sessions improved my study skills | 1,806 | 2.79 | 0.82 |

*Note.* 1 = *strongly disagree*, 2 = *disagree somewhat*, 3 = *agree somewhat*, and 4 = *strongly agree*.

Most respondents found the student study sessions to be useful (see Figure 71).

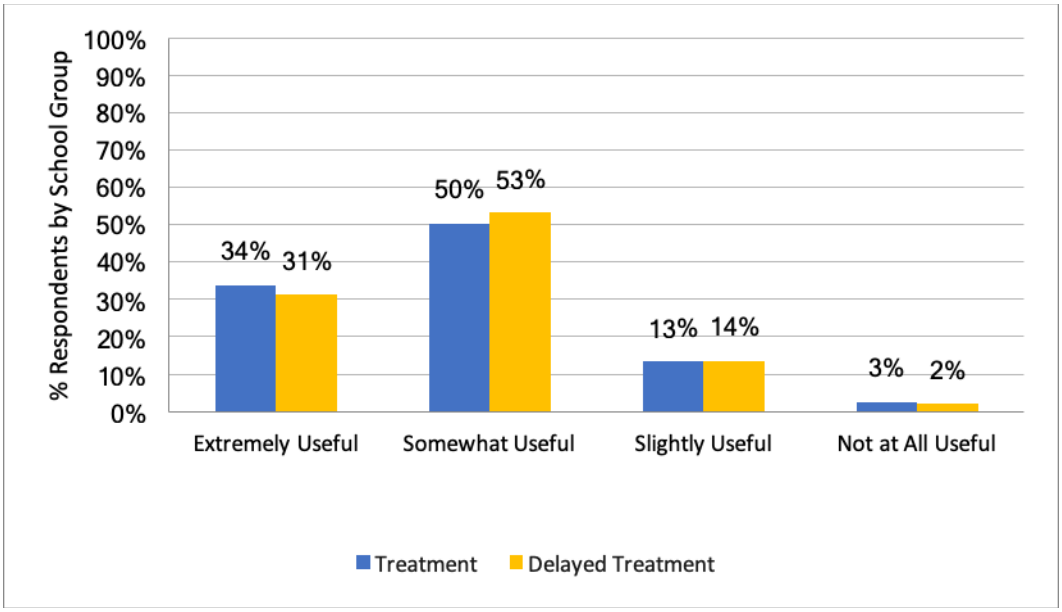



Figure 71. Qualitative assessment of student study sessions.

### 6. 2018–2019 Student Survey: AP Preparation

Survey participants have a high degree of confidence in their teachers' content knowledge (see Table 99).

Table 99

*Confidence Metrics*

| Description of student study session | *n* | *M* | *SD* |
|---|---|---|---|
| My AP teacher understands the content being taught | 2,691 | 3.79 | 0.51 |
| I am confident in my ability to successfully complete AP courses | 2,689 | 3.33 | 0.70 |
| I am confident in my ability to learn new STEM content | 2,689 | 3.30 | 0.68 |
| I am confident in my ability to get a score of 3 or better on the AP exam | 2,687 | 3.07 | 0.79 |
| I was nervous about how hard the AP course would be when I signed up for it | 2,687 | 3.06 | 0.93 |

*Note.* 1 = *strongly disagree*, 2 = *disagree somewhat*, 3 = *agree somewhat*, and 4 = *strongly agree*.

### 1. 2018–2019 Student Survey: AP Exams

The most cited reason for not taking the AP exam for the course in which the student completed the survey was not feeling ready (see Figure 72).

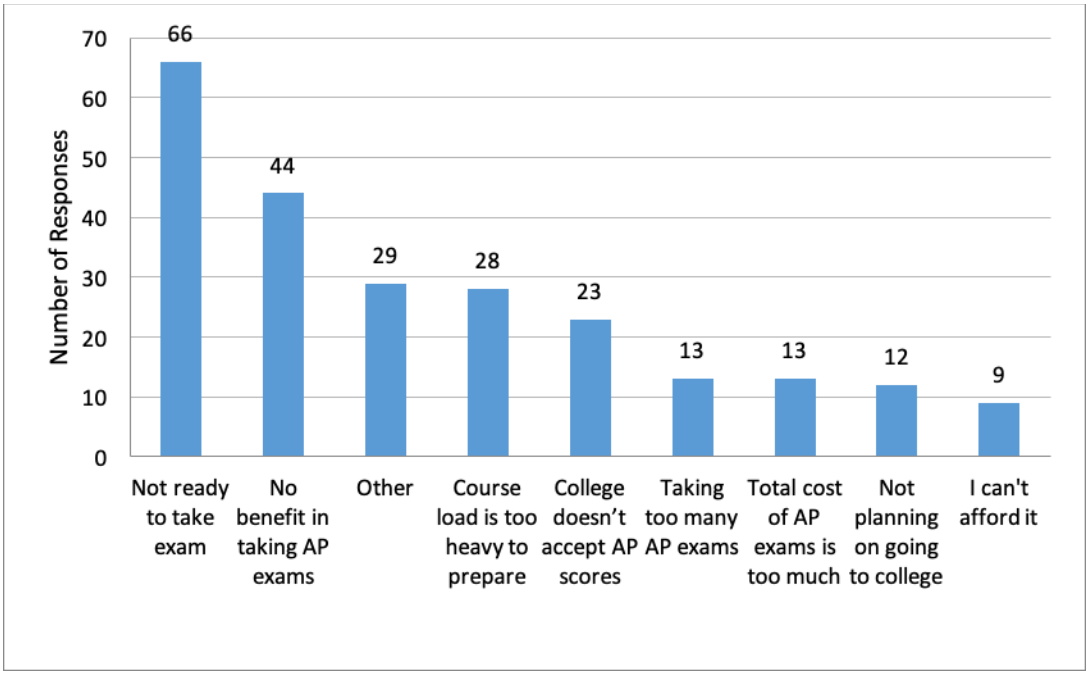



Figure 72. If Respondent is *not* taking the exam, reasons for decision ($n$ = 165) (multiple selection).

### 1. 2018–2019 Student Survey: Student Incentives and Rewards

In addition to the cash reward from NMSI, many schools offer other incentives to participate in AP courses such as weighted GPAs. Figure 73 illustrates the incentives offered at Program Schools.

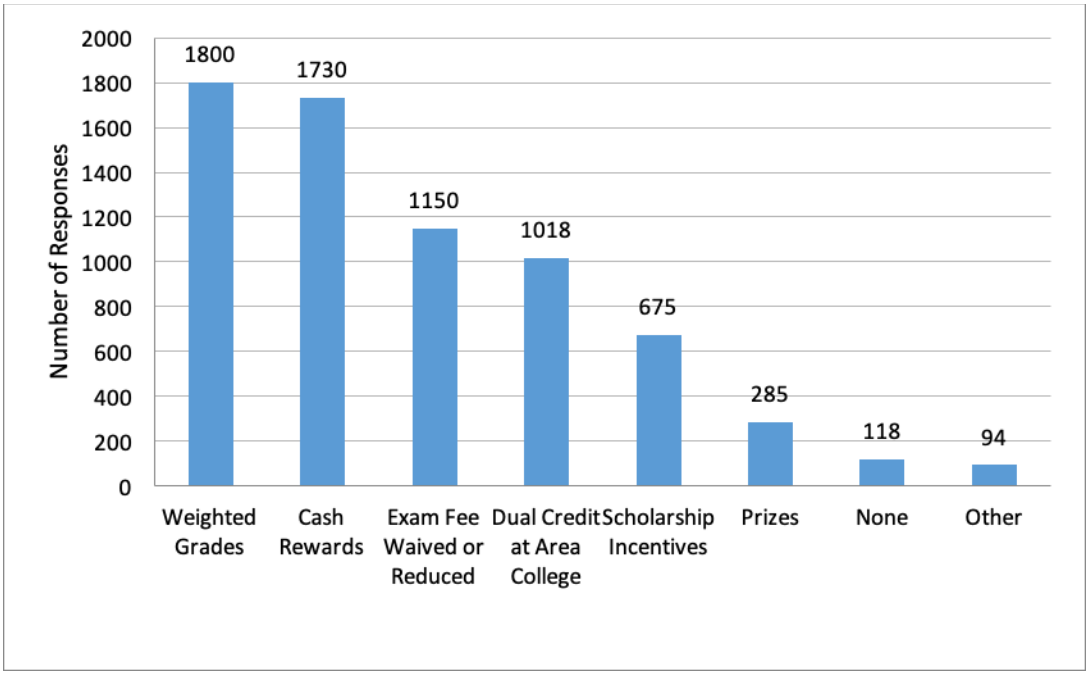



Figure 73. AP course incentives (*n* = 2,664) (multiple selection).

The majority of survey participants (57%) said that the cash awards were at least somewhat important in encouraging their participation in AP courses (see Figure 74).

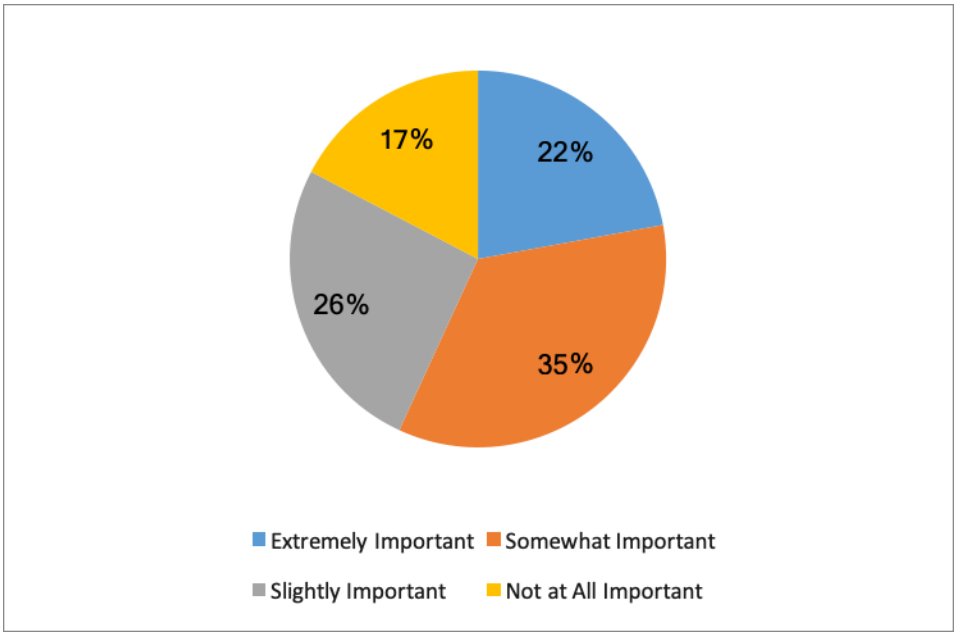



Figure 74. Importance of cash awards in encouraging AP course participation (*n* = 2,310).

## VI. Summary

A summary of findings from the teacher and administrator surveys and interviews follows.

In 2018, data were collected from 200 teachers, 62 administrators, and 1,930 students. Data were collected in 2019 from 211 teachers, 55 administrators, and 2,710 students. Implementation information was collected from administrative records, surveys, and interviews.

As in previous years, teacher training was seen as the most effective component of the CRP and mentoring the least effective. Teachers reported positive feedback on the training and indicated an increase in content knowledge and training on instructional strategies and techniques. This suggests the CRP PD is meeting a need and helping prepare teachers to feel better equipped to teach AP courses.

*Study sessions*: A major component of the CRP is providing students with 12 hours of instruction outside of their normal classroom experience. Conducted on three Saturdays throughout the school year, Student Study Sessions are held at a local high school and include students from all of the participating schools in the area. Students rotate to different classrooms over the course of four hours, receiving instruction from a number of seasoned educators with different experiences and expertise from their usual instructor. Saturday instruction provides many challenges, and schools employee innovative methods of incentivizing and facilitating attendance such as offering extra credit and providing students with transit passes. Students were asked how they learned about the Student Study Sessions, how many they attended, in what ways the schools incentivized or facilitated their attendance, how they could make up for a missed session, and finally how useful the Student Study Session component of the CRP was.

Student study sessions were viewed as an important and positive component of the program. The AP courses cover a lot of content and the study sessions were seen as helpful for providing students' additional instructional time as well as teaching support and guidance for teachers. Many teachers mentioned in the interviews that the study sessions provided excellent resources for classroom use. Attendance at the study sessions was not as high as we would have expected, and one reason for this may be the scheduling issues. Students who responded to the survey indicated lower levels of agreement on the convenience of study session scheduling. Saturday scheduling posed significant problems for many students because of work and sports team commitments as well as other family-related obligations.

Surveys were returned from both treatment and delayed treatment Schools and so we had the opportunity to explore differences between them. In 2019 gaps narrowed between Treatment School and Delayed Treatment School support of Student Study Sessions in only two of three metrics (providing transportation, implementing award systems, and enabling make-up

opportunities). Despite an increase in the number of students taking one or two AP courses, thus reducing the number of Saturday sessions, self-reported attendance in Delayed Treatment Schools declined from Year 2 to Year 3. Student assessment of the usefulness of the Student Study Sessions was unchanged for Treatment Schools from year to year, but the Delayed Treatment School evaluation was more positive in Year 3 (84% "somewhat" or "extremely" useful).

*Teacher incentives*: While teachers consistently said incentives were not an important motivating factor, they often clearly thought of them as (poor) compensation for the many additional hours of work involved in the program—particularly training and the student study sessions. Many teachers also reported that they would be teaching the AP courses regardless of the incentives—indeed some had been teaching the courses for several years. Teachers did, however, indicate that the incentives were nice to have and a welcome added bonus for their participation.

*Training*: If the number of students enrolling in AP courses continues to increase, there may be a need for more training in differentiated instruction (for students of varying ability levels). This is an area (helping them differentiate instruction for students at different ability levels) in which teachers tend to find the training sessions less effective than other areas.

*School culture*: In 2018, 64% percent of teachers and 77% of administrators felt that school culture had changed (in a positive way) since the implementation of the CRP. In 2018, when comparing the views of delayed treatment group teachers (in year one of the program) to treatment group teachers (in year two), we found those with more experience with the program, believed more strongly that the CRP contributed to improvements in student experience with STEM AP courses, teacher content knowledge, teacher instructional skills and strategies, the school culture of continuous improvement, and school leadership valuing STEM learning. Some schools in the sample indicated they had already been encouraging students to take AP courses and so the culture shift or improvement would not likely be as marked as in schools with no existing AP program. In 2019, when we had teachers with between one and three years of experience with the CRP, we found that survey participants felt that their schools' administration "promoted a culture of continuous improvement" and "valued STEM learning" at a higher rate than they felt the school set clear goals for either AP enrollment or exam performance.

Sixty-two teachers completed surveys for each of the three study years. Results indicated:

- Teachers saw the spring and fall trainings as having less value than the summer training.

- There was a significant decline in the number of students attending all three study sessions.
- Teachers' ratings of the quality and impact of the study sessions increased, although their level of satisfaction relating to convenience decreased.
- Over time, incentives were seen as less important.
- Over time, teachers had an increased use of NMSI resources.
- Across the years, teacher assessment of their own understanding of AP course concepts increased.
- There was a steady increase in belief that the CRP is an effective way to increase student enrollment in AP courses.

Below we briefly summarize the key findings from the student survey measure:

- Most students learned about AP and the CRP from their teachers or school counselors, and more than a third of students learned about the AP courses from other students.
- Across both years, most student respondents indicated they planned to attend some sort of postsecondary institution.
- Typically, (across both years) around 75% percent of students attended at least one student study session for the course during which they completed the survey, and approximately 20% attended all of the study sessions available for that course. On average students attended 1.55 study sessions. Around 80% of students found the study sessions to be extremely or somewhat useful. On average, students had the lowest level of agreement that the study sessions were conveniently scheduled, which sheds additional light on the issue of poor attendance.
- Students indicated their AP teachers had good content knowledge, and on average agreed with statements relating confidence in their ability to achieve a qualifying score, learning STEM content, and taking AP courses and exams.
- Most students surveyed indicated they were planning on taking the AP exam. For the ones who were not taking it (in 2018 $N = 98$; and in 2019 $N = 165$) the reason given was they did not 'feel ready to take the exam'. and two thirds indicated they were offered cash incentives for performance on AP exams.
- Over half of students rated the financial incentives as extremely or somewhat important in encouraging their participation in AP courses.

## A. Discussion

The objective of the CRP is to increase academic intensity and access to rigorous courses, improve student achievement, and decrease the college readiness gap, especially among traditionally underrepresented students. Our evaluation of the CRP consisted of three parts: (1) measuring the program's impact on selected student AP exam outcomes, (2) determining the

impact of the program on school perspectives and culture, and (3) assessing of the fidelity of implementation of the CRP at the school level.

As discussed above, program impact was evaluated using a 2-level hierarchical generalized linear model (HGLM) with students nested within schools. AP exam data from 48 Treatment Schools, with a total of 8,778 exams in 2018 and 9,378 in 2019, and 48 matched control schools, with 7,505 exams in 2018 and 2019 in Year 3, were analyzed for this study. First a non-conditional model was use, and then a conditional model in which school size was considered. One of the goals of the CRP is to increase acess to AP courses—particularly in STEM. This is clearly working. In 2018 the probability of a student taking an AP exam in the Program Schools was, on average, 7% higher than the paired Comparison Schools, and the difference was statistically significant. And in 2019 the effect was even greater with the probability of taking an AP exam being significantly higher in Program Schools (18%) than in Comparison Schools (3%). Although a smaller effect, when looking at the probability of an exam yielding a qualifying score, while there was no significant difference between the two groups in 2018, in 2019, however, exams taken at the Program Schools had a significantly higher overall probability (2%) of receiving a qualifying score than the Comparison Schools (0%). These analyses compared results to the total school population. In the next analyses, we looked only at those students who took AP exams. In 2018, overall, the fitted probability of achieving a qualifying score among the exams taken was 8% in the Program Schools, compared to 22% in the Comparison Schools. In 2019, however, the difference between the Program Schools (7%) and the Comparison Schools (9%) was not statistically significant.

Given differences in school sizes program impact was further evaluated using a conditional HGLM. Findings were similar as to those from the unconditional model for the number of students taking AP exams, and for the number of qualifying scores (when looking at the proportion of school population). When only looking at exam results for those who took the exam, the fitted probability of receiving a qualifying score among the exams taken was again higher in the Comparison Schools (as was the case in the unconditional model results). However, the difference between the two groups was not significant with a gap narrowed down from 14% to 10%.

One of the CRP goals is to increase enrollment in AP courses—particularly for students who may not typically see themselves as “AP students.” We saw higher levels of student enrollment in the Program schools, thus more students were exposed to academic courses in which they engaged in college-level work which can help increase college aspirations and identity. While there was a higher percentage of qualifying scores at the control schools, this may be because not all students in those schools are required to take the AP exam if they take an

AP course (which was a CRP requirement). This could result in only those students who felt confident they would pass the AP test, actually taking the test in the control schools.

Heterogeneity in program impacts was explored using some of the fidelity matrix elements including implementation 'dosage' (two or three years), years of teaching experience, AP-specific teaching experience, PSD's level of motivation for the CRP, open enrollment, beliefs about CRP effectiveness, and teacher engagement. School means for all of these measures were used. AP teaching experience was positively and significantly related to all outcomes in the impact study and all matrix elements mentioned were related to an increased probability of achieving a qualifying score. In addition, in 2018, students in schools with two years of implementation were more likely (36%) to take an AP exam than students in schools with one year of the CRP (13%). In 2019, however, the difference was not significant.

Data collected for the fidelity matrix indicated that not all elements of the CRP were implemented with high fidelity; In 2018, results indicated that 43 out of 48 schools (90%) achieved 80% or better implementation fidelity, for an average fidelity score of 89%. Four schools achieved a perfect 100% fidelity score. In 2019, 88% of schools achived 80% or better implementation fidelity. Ten schools achieved a perfect 100% fidelity score. In 2018 in more than 75% of schools, not all teachers fulfilled their requirements for attending the required teacher training sessions, and so this component was not implemented with fidelity. In 2019, this picture improved a little with 15 schools (31%) meeting the 80% threshold. Teacher stipends, administrator bonuses, and student qualifying score awards were paid as expected.

Teacher survey data indicated teacher training and other professional development activities were seen as the most effective components of the CRP and mentoring the least effective. Over half of teachers expressed interest in online collaboration and networking with AP teachers and other AP experts, something which could be implemented in lieu of the under-utilized mentoring component.

Teachers reported positive feedback on the training and indicated an increase in content knowledge. This suggests the CRP PD is meeting a need and helping prepare teachers to feel better equipped to teach AP courses. Only 16% of teachers felt that student financial incentives were either the most important or second most important CRP component. Teachers did, however, view the student incentives as an important program component to encourage student enrollment in AP courses. Likewise, students rated the financial incentives on average as somewhat to extremely important in encouraging them to participate in AP courses.

Student study sessions were viewed as important and positive components of the program. The AP courses cover a lot of content and the study sessions were seen as helpful for providing

students' additional instructional time as well as additional pedagogical support and guidance for teachers. Many teachers mentioned that while the study sessions provided excellent resources for classroom use and were beneficial to students, some of the students who would benefit from the sessions were unable to attend. An issue which is amplified for students taking multiple AP courses. These findings suggest that it may be beneficial to rethink the scheduling and location of the student study sessions. makes ensuring adequate and accurate data on student attendance an important goal for moving forward, as well as discussing possible alternatives to attending these sessions at the weekends, in person.

Taken as a whole, the individual components of the CRP aim to push school cultures toward greater inclusion, higher expectations, and an emphasis on STEM education. Survey participants felt that their schools' administration "promoted a culture of continuous improvement" and "valued STEM learning" at a higher rate than they felt the school set clear goals for either AP enrollment or exam performance which may be an opportunity for clearer messaging and goal setting in the future. In addition, teachers and administrators in some schools in the sample indicated they had already been encouraging students to take AP courses (prior to implementation of the CRP) and so the culture shift or improvement would not likely be as marked as in schools with no prior AP program.

Evaluation of the CRP from the teacher and administrator perspective provides both unique insight as well as valuable support for the fidelity of implementation data we are gathering as part of our evaluation study. Perspectives of those on the ground who are implementing the CRP helps determine which factors are most important in creating and sustaining an accessible and successful AP program. Survey and interview data from CRP teachers across the country provided valuable support for data collected for the implementation evaluation and helped determine how CRP participants perceive the importance and efficacy of key program elements.

We conducted our evaluation over three years of NMSI program implementation, during which teachers continued to receive progressively more rigorous training and lessons; teachers and administrators continued to push further toward increasingly challenging goals; and both students and teachers receive nominal monetary awards for success. We saw many instances of teachers with two or three years of experience in the program attributing improvements in their own teaching, student experiences with AP STEM courses, shifts in school culture toward improving access to AP courses (in particular the impact of open enrollment), and the development of a culture of continuous improvement to NMSI's CRP.

## VII.References


ACT. (2016). *The condition of college & career readiness 2016.* http://www.act.org/content/dam/act/unsecured/documents/CCCR_National_2016.pdf

Chen, X. (2013). *STEM attrition: College students' paths into and out of STEM fields* (NCES 2014-001). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Cross, I., Kilpatrick, C., & LaMonica, L. (2012). *Evaluation of the Advanced Placement Training and Incentives Program: Teachers' perspectives.* Thomas Jefferson Program in Public Policy at the College of William & Mary.

Fleischman, H. L., Hopstock, P. J., Pelczar, M. P., & Shelley, B. E. (2010). *Highlights from PISA 2009: Performance of US 15-year-old students in reading, mathematics, and science literacy in an international context* (NCES 2011-004). National Center for Education Statistics.

Kena, G., Aud, S., Johnson, F., Wang, X., Zhang, J., Rathbun, A., Wilkinson-Flicker, S., and Kristapovich, P. (2014). *The Condition of Education 2014* (NCES 2014-083). U.S. Department of Education, National Center for Education Statistics. http://nces.ed.gov/pubsearch

Matthews, J. (2018, December 7). A test of critical thinking: Why don't all AP students take AP tests? *Washington Post.*

National Science Board. (2010). *Science and engineering indicators 2010* (NSB 10-01). National Science Foundation.

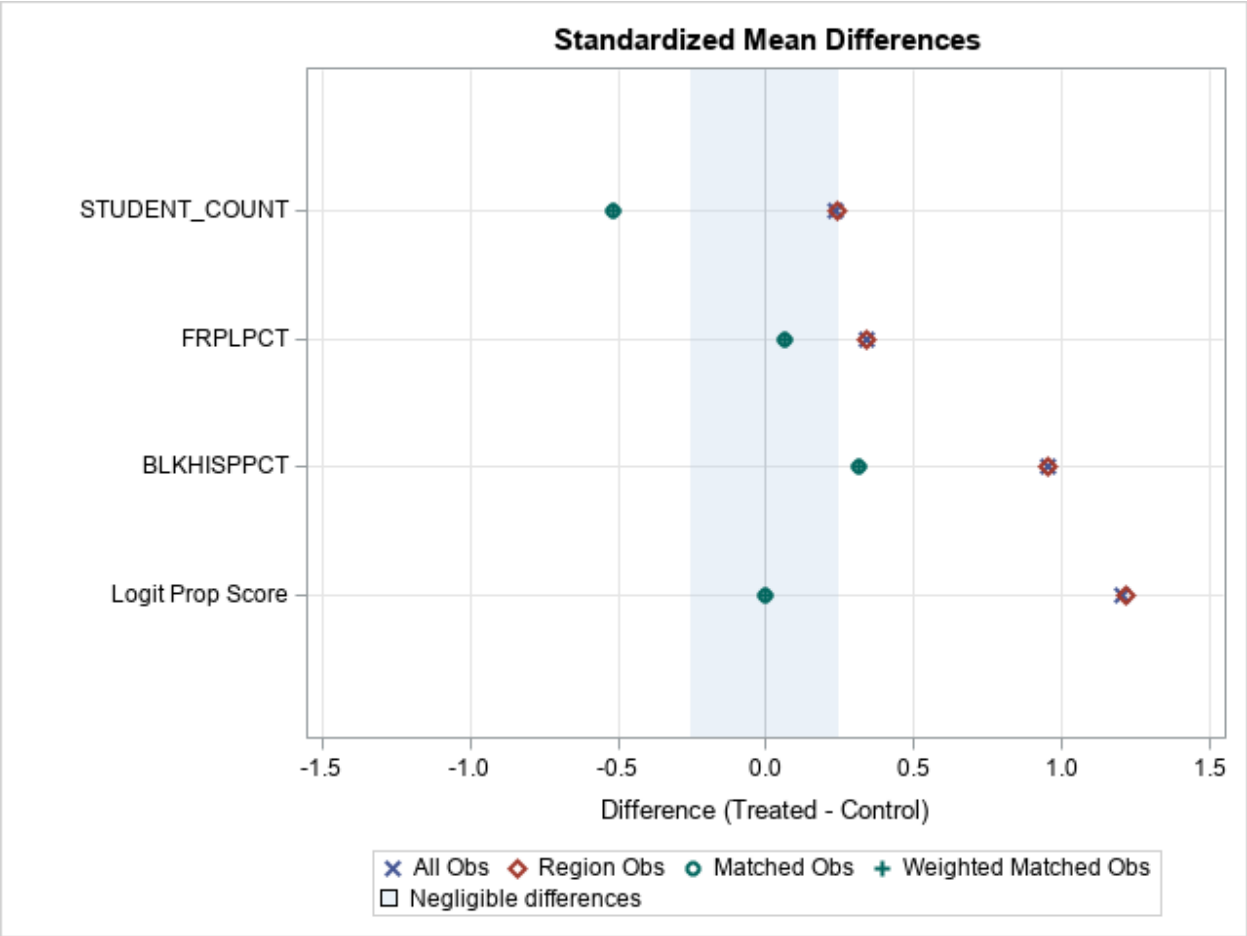
National Science Board. (2014). *Science and engineering indicators 2014* (NSB 14-01). National Science Foundation. https://www.nsf.gov/statistics/seind14/content/etc/nsb1401.pdf

Pew Research Center. (2017). *U.S. students' academic achievement still lags that of their peers in many other countries.* http://www.pewresearch.org/fact-tank/2017/02/15/u-s-students-internationally-math-science/

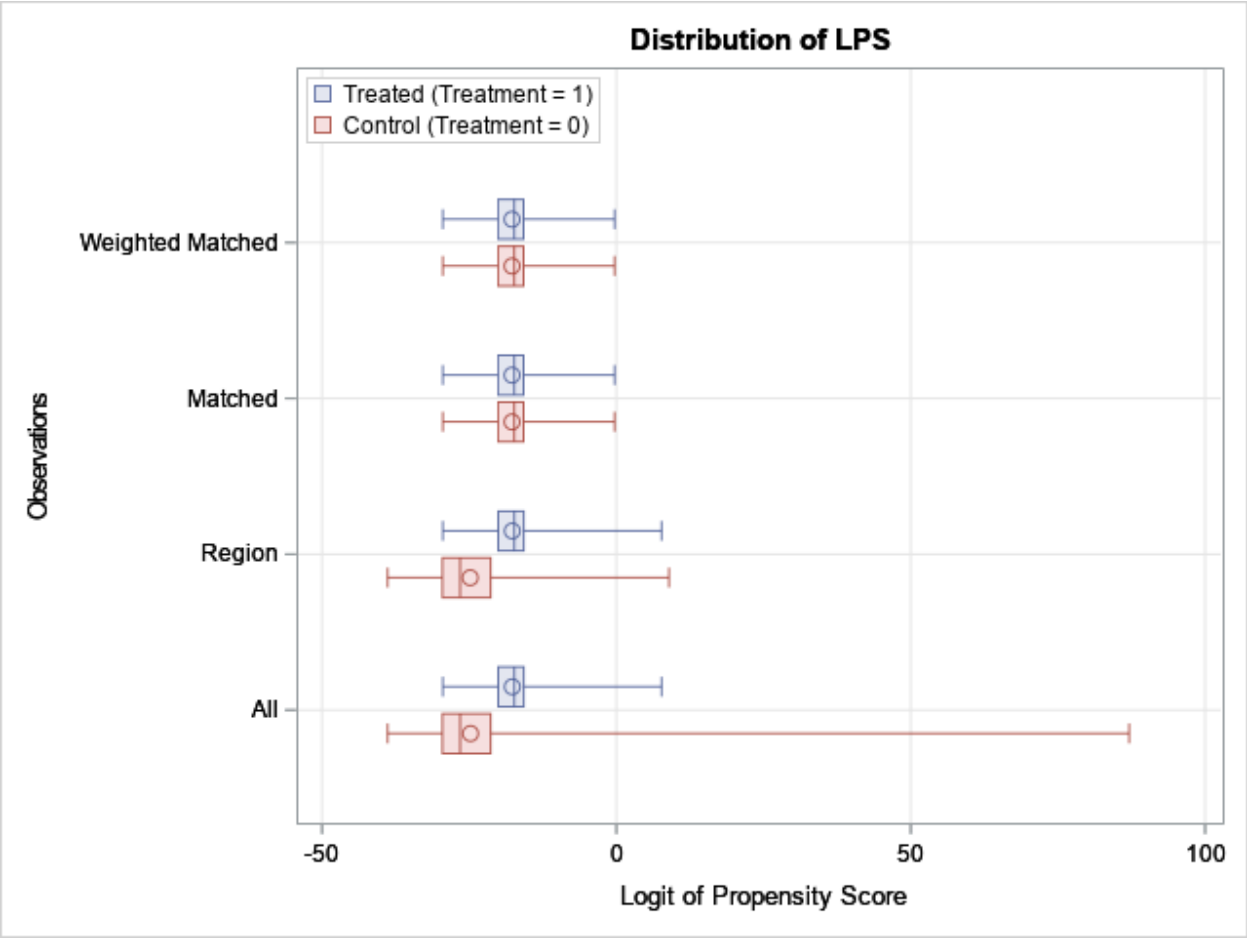
Sherman, D., Darwin, M.J., Song, M., Li, Y., & Stachel, S. (2015, April 17). *First-year impacts of the National Math and Science Initiative's Advanced Placement Training and Incentive Program on high school students' outcomes*. Paper presented at the 2015 annual meeting of the American Educational Research Association. Retrieved from the AERA Online Paper Repository.

## Appendix A: Propensity Score Matching Results

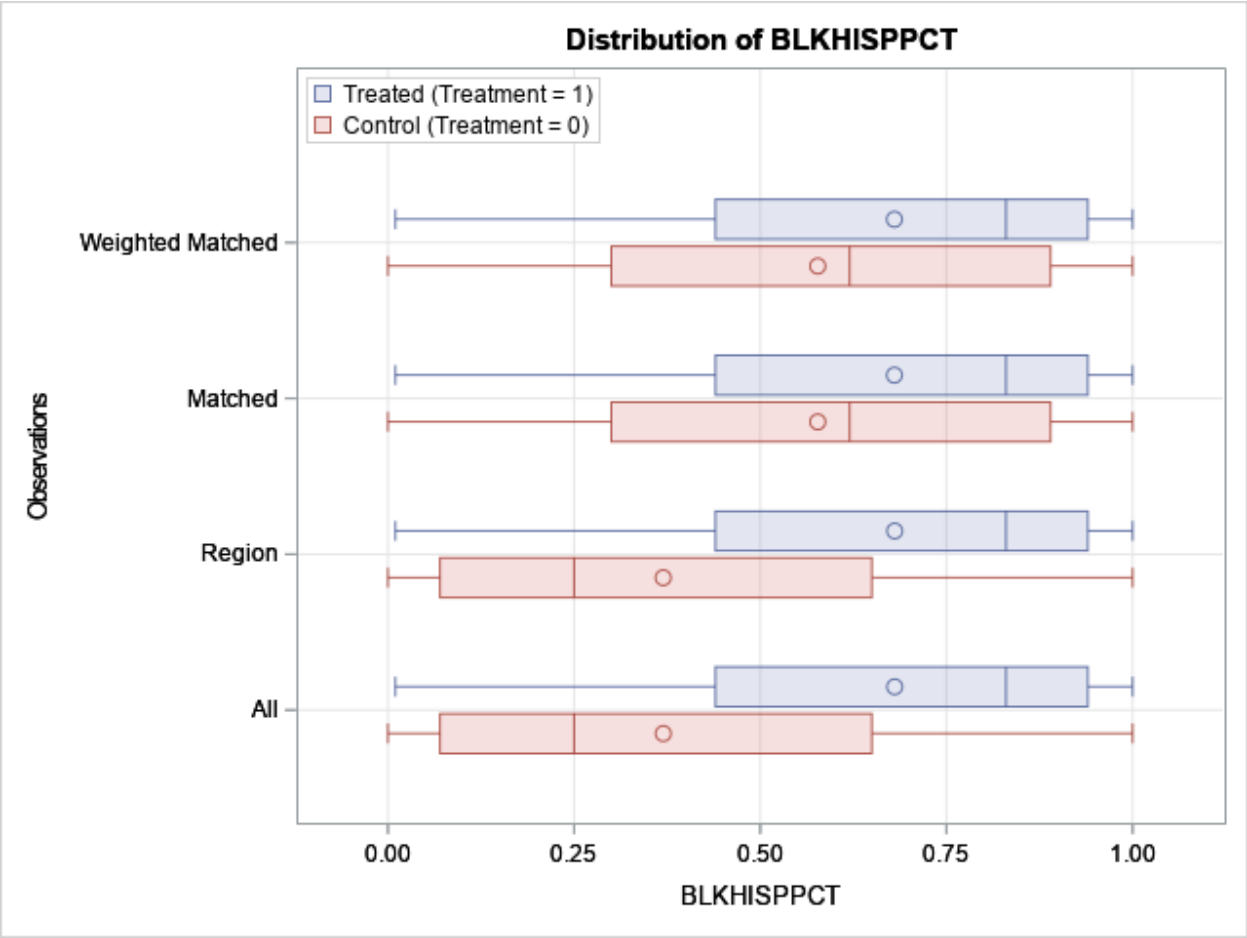
A. Standardized mean differences



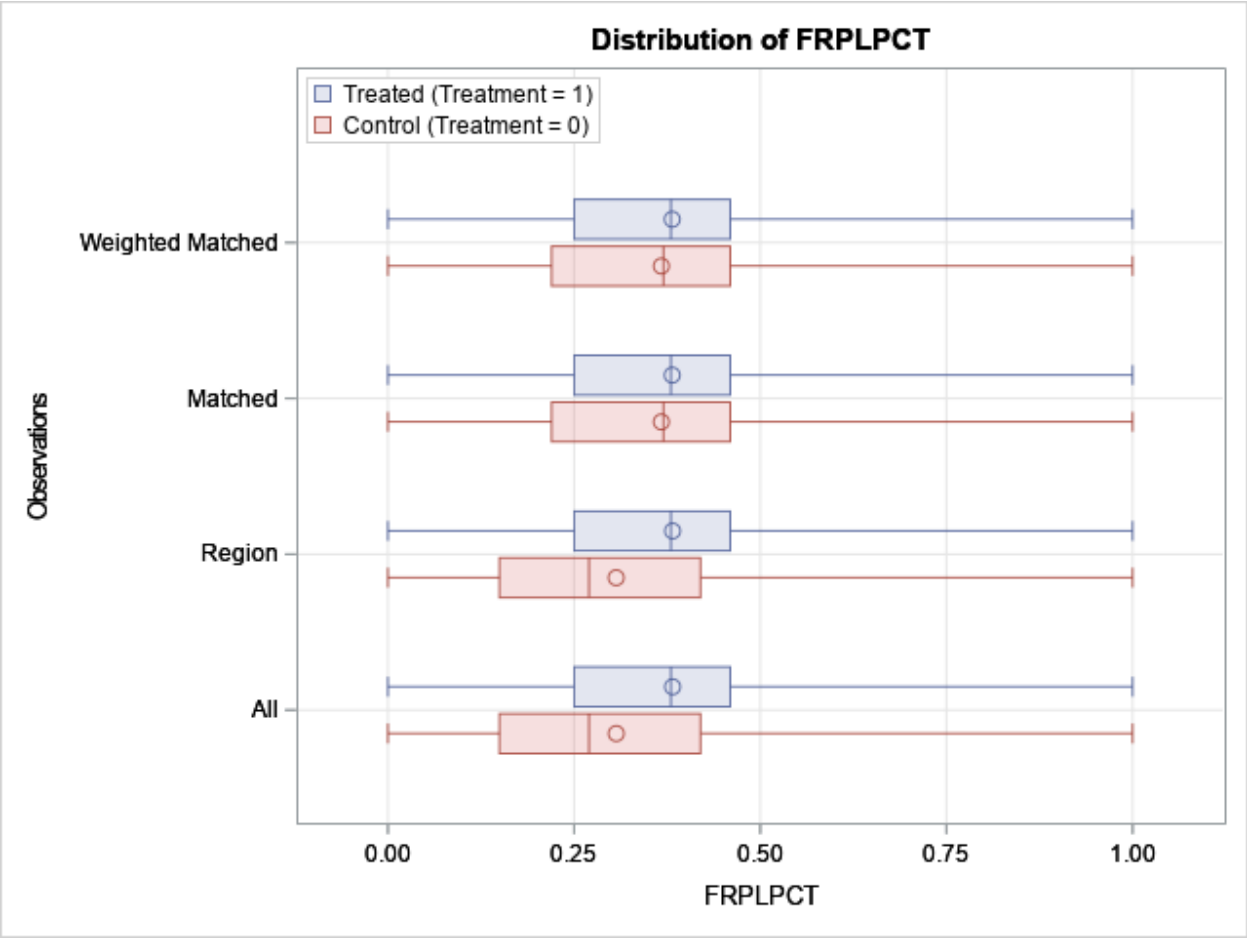
B. Distribution of logit of propensity score



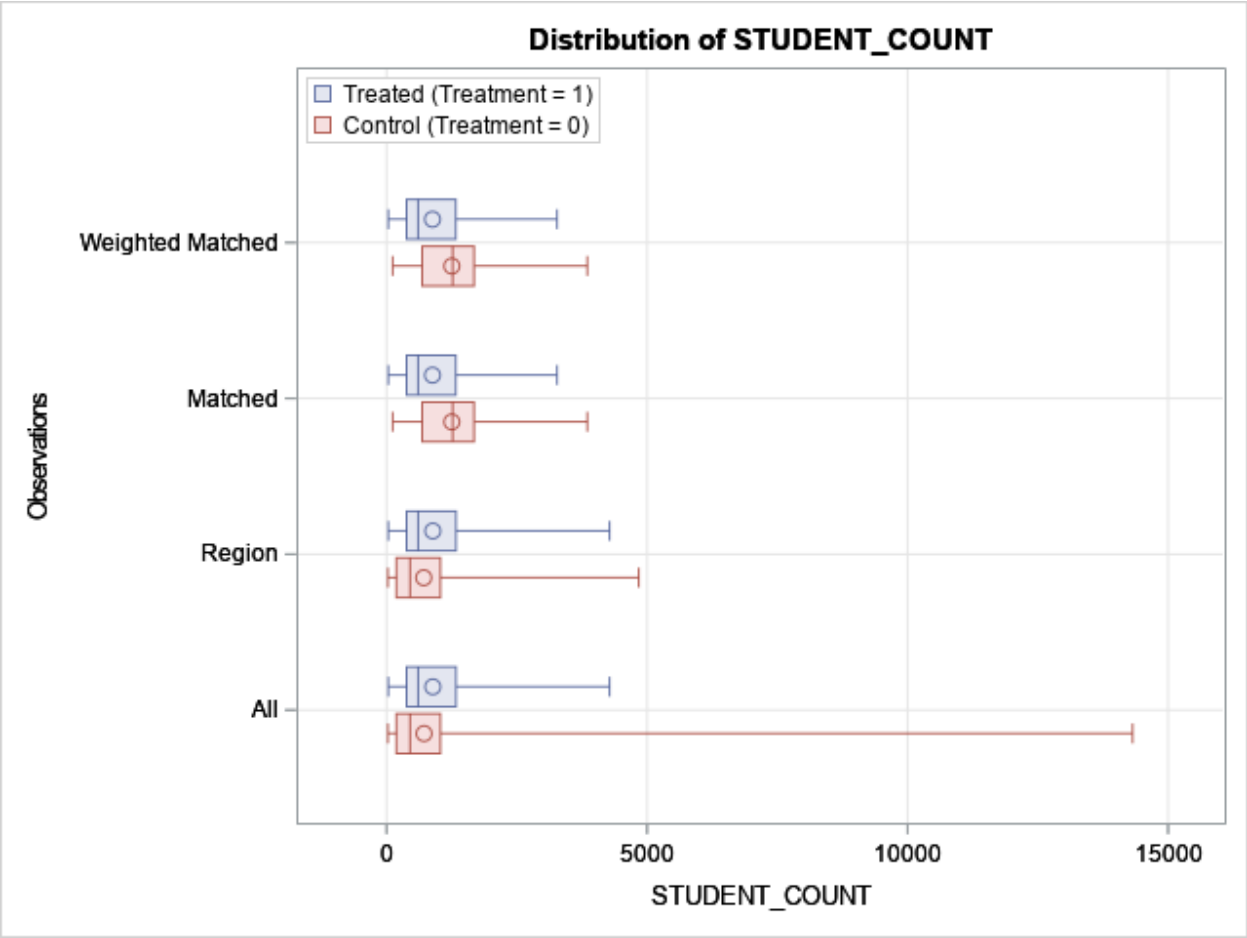
C. Distribution of the percentage of Black and Hispanic students



D. Distribution of the percentage of students eligible for free/reduced priced lunch



E. Distribution of enrolled number of students



## Appendix B: Hierarchical Generalized Linear Model

We used a 2-level hierarchical generalized linear model (HGLM) with two different types of cluster, school and matched pair. We fist present Standard HGLM analysis with school id as clusters, followed by Matched pair analysis using matched pair of NMSI and comparison schools as clusters. In each set of analyses, three subject domains, English, Math and Science, were examined using AP exam data (2017–2018 and 2018–2019).

The outcomes of interest for the 2-level HGLM were binary: (1) Total tests, whether or not a student took at least one AP exam (coded as 0 for not taking an AP exam; 1 for taking one or more AP exam), and (2) Score 3 or higher, whether or not a student earned a qualifying score of 3 or higher on at least one AP exam in a subject category (coded as 0 for not achieving qualifying scores in AP exams; 1 for achieving qualifying scores in AP exams); (3) Score 3 or higher among exam takers, the second outcome with a restricted sample of exam takers.

In each of the Standard HGLM analysis and the Matched pair analysis, unconditional and conditional HGLMS were fitted. For the unconditional models, only treatment variable was included in the analysis models, while school size was added as a covariate to increase the precision and control for the possible confounding in the conditional models.

For the Standard HGLM analysis, the response of student *i* in school *j*, $Y_{ij}$, was binary, which was coded as 0 or 1. For each outcome, the probability of response equals to one, $p_{ij} = P(Y_{ij} = 1)$ was modelled using a logit link function, $\eta_{ij}$. In this way, the Level-1 model is specified as follows.

Level 1:

$$\eta_{ij} = log\left(\frac{p_{ij}}{1 - p_{ij}}\right) = \beta_{0j}$$

In the above equation, $p_{ij}$ is the probability of student *i* in school *j* to take at least one AP exam, score 3 or above in the AP exam or score 3 or above in the AP exam among exam takers. We assume that the average log odds of school *j* varies across schools, represented by the random intercept $\beta_{0j}$.

The corresponding Level-2 model is expanded to include the treatment variable which indicates whether school *j* is a NMSI school or a comparison school as well as the number of students in school *j* ($Size_j$) as a covariate.

Level 2:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}Trt_j + \gamma_{02}Size_j + u_{0j}$$

In the School-level model, $\gamma_{00}$ represents the overall average log odds for comparison schools with average school size. The coefficient for the treatment indicator, $\gamma_{01}$, is the difference in log odds between treatment and comparison schools, controlling for school size. The coefficient for school size, $\gamma_{02}$, means the change in the log odds associated with one unit change in school size. The random error, $u_{0j}$, is the deviation of school j from the predicted log odds given its treatment status and school size.

For Matched pair analysis, the probability of the response of student *i* in pair *k*, $Y_{ik}$, equals to one was modelled with a logit link function, $\eta_{ik}$, similar with the Standard HGLM analysis. The Level-1 model is expressed as follows.

Level 1:

$$\eta_{ik} = log\left(\frac{p_{ik}}{1-p_{ik}}\right) = \beta_{0k} + \beta_{1k}Trt_{ik}$$

Unlike in Standard HGLM analysis, $p_{ik}$ now represents the probability of student i in matched pair k to take at least one AP exam, receive a score 3 or above in the AP exam or score 3 or above in the AP exam when the sample is restricted to the exam takers. The random intercept, $\beta_{0k}$, is the average log odds for comparison group in pair k while the fixed slope, $\beta_{1k}$, is the expected difference in log odds between treatment and comparison group. Note that the treatment indicator for student i in pair k, $Trt_{ik}$, is treated as a Level-1 variable.

The Level-2 model with a random intercept and a fixed slope is written as follows.

Level 2:

$$\beta_{0k} = \gamma_{00} + \gamma_{01}Size_k + u_{0k}$$

$$\beta_{1k} = \gamma_{10}$$

The interpretation of each coefficient from Matched pair analysis is similar with Standard HLM analysis. In the above equation, $\gamma_{00}$ is the overall average log odds for comparison group, controlling for the covariate $Size_k$, which is computed as the mean of the number of students in pair j. $\gamma_{01}$ is the change in the log odds associated with one unit increase in $Size_k$. The random error of pair k, $u_{0k}$, now represents the deviation of pair k from the predicted log odds given the covariate, $Size_k$. The treatment slope, $\beta_{1k}$, is assumed as fixed to have the same value, $\gamma_{10}$, across all the pairs, which indicates that the difference in log odds between treatment and comparison group is the same in all pairs.

As the logit link function is used for binary outcomes, we exponentiated all the coefficients when interpreting the results in probability scale instead of log odds. For the Standard HGLM analysis results, the expected probability of having the response equals to one for Treatment

Schools is $\hat{p}_{trt} = \frac{\exp(\hat{\gamma}_{00}+\hat{\gamma}_{01})}{1+\exp(\hat{\gamma}_{00}+\hat{\gamma}_{01})}$, and $\hat{p}_{comparison} = \frac{\exp(\hat{\gamma}_{00})}{1+\exp(\hat{\gamma}_{00})}$ for comparison schools. For the Matched pair analysis, the expected probability for NMSI schools was $\hat{p}_{trt} = \frac{\exp(\hat{\gamma}_{00}+\hat{\gamma}_{10})}{1+\exp(\hat{\gamma}_{00}+\hat{\gamma}_{10})}$, and $\hat{p}_{comparison} = \frac{\exp(\hat{\gamma}_{00})}{1+\exp(\hat{\gamma}_{00})}$ for comparison schools.

## Appendix C: 2019 Teacher Survey Training Changes from SY2017–2018

| Subject Taught | Comment about Changes to Training from the 2017–2018 School Year |
|---|---|
| English Language and Composition | The AP Lang resources in 18-19 seemed of a poorer quality than those used in 17-18. The summer and fall instructor I had (she was the same for both) did not provide me with a high-quality learning experience. I left both PD experiences disappointed. The spring session, however, was excellent. The facilitator was much more skillful at leading adult PD. |
| English Literature and Composition | Less NMSI people at events |
| English Literature and Composition | The entire process of booking travel has been more streamlined and user-friendly |
| English Lit (and English Lang) and Composition | I think my trainings were LESS impactful, with the exception of the summer APSI. |
| English Language and Composition | some of the materials were disappointing for the APSI and spring training |
| English Literature and Composition | I think the teacher trainings have gotten weaker due to the lack of materials/content being offered for AP Lit. |

## Appendix D: 2019 Teacher Survey Financial Awards Changes from SY2017–2018

| School Group | Teacher Year in Program | Changes to Financial Awards from 2017–2018 School Year |
|---|---|---|
| Delayed Treatment | 2 | The scheduling of Student sessions at the same time as Teacher trainings were held meant that not all stipends could be received. It does not feel fair that the money is offered but then scheduling prevents one from receiving it. |
| Treatment | 2 | The % of students who must earn a 3 or higher for me to receive my incentive increased substantially. |
| Delayed Treatment | 2 | Not paid for attending spring training. The training was during our spring break and no incentive was awarded to attend the session during our spring break. |
| Delayed Treatment | 2 | Awards became significantly more difficult to achieve (not attainable) |
| Delayed Treatment | 2 | goals were set that were not attainable due to schedule and room size |
| Treatment | 3 | Increased my target goal to receive bonus |
| Delayed Treatment | 2 | Teachers goals in some cases were not attainable |

## Appendix E: 2019 Teacher Survey Preferences for Additional Tools and Materials

| Subject Taught | Comments about Preferred Additional Tools and Materials |
|---|---|
| English Language and Composition | the materials ordered that I never received. |
| Calculus & Statistics | Test banks |
| Chemistry | short and content specific hands on |
| English Literature and Composition | I *REALLY* need help with how to SCORE. |
| Biology | Materials that are post re-write of the AP exam and more similar/aligned to the AP exam |
| Chemistry | Materials that are all current with the revised AP Chem Exam |
| Physics | More useful Saturday Sessions. The instructors just talk and the kids just sit there. |
| English Literature and Composition | BOOKS! I want to be able to provide each of my AP students with high-interest, rigorous texts to read and analyze. |
| Biology | I am satisfied with the current offerings |
| Chemistry | More multiple choice items for some of the units where there aren't many beyond those in the unit exams. |
| Calculus & Statistics | Pre made exams |
| Physics | Post all previous released exams |
| Calculus | I believe we have a plethora of resources at our fingertips |

## Appendix F: 2019 Teacher Survey CRP Impact on AP Enrollment

| School Year in Program | Subject Taught | Comment about Impact on AP Enrollment |
|---|---|---|
| Delayed Treatment | Calculus & Statistics | It can be |
| Delayed Treatment | English Language and Composition | I feel that AP attendance rates stayed relatively the same for AP classes with or without NMSI |
| Treatment | Biology | Our administrator did very very little. He should not cut and paste information into spreadsheets himself. The Lady who ran Saturday's was much more competent. |
| Treatment | Environmental Science | Our original School Director left unexpectedly. The person who picked it up as been doing an excellent job but has been learning about the program as he goes. |
| Delayed Treatment | Computer Science | Know its useful |
| Delayed Treatment | Calculus | I wish that I could be in charge of my students study sessions and receive a stipend instead of the CRP Manager at our school |
| Delayed Treatment | English Literature and Composition | I felt this program was extremely beneficial to my students and me. |
| Delayed Treatment | Physics | I feel there is still a need for additional supports if more students are taking AP courses. Supports such as time, and more scaffolded materials. |
| Treatment | English Literature and Composition | I don't know who our CRP Program Manager is this year. |
| Treatment | Biology | I would like to have more! This is the end of the 3 years, and everyone--including kids want to continue. |
| Treatment | Calculus | My school had a very high AP enrollment before the CRP program. |
| Treatment | Chemistry | We had a change in leadership. |
| Delayed Treatment | Physics | poor communication |
| Treatment | Statistics | Our's switched mid year from Camarena to Anderson (both were good but the transition was slow) |
| Treatment | English Literature and Composition | I don't think our Partner School Director does a very effective job of communicating between teachers and NMSI. I often find out important information from other teachers in the program (who have found out the information from their Partner School Director) |
| Treatment | Physics | Effective:yes. Efficient: No. |
| Treatment | Physics | I think the CRP gave the district a push to look at students who normally would not sign up and change our culture but I do not think that the CRP itself would be the end all for that. That same result could have come from administration. |

## Appendix G: 2019 PSD Survey Comments about CRP Matching Expectations

### More Details on Why CRP Did or Did Not Meet Expectations

- I received support after I asked but was new to the program and "didn't know what I didn't know".
- This year it wasn't as structured as previous years.
- It is impossible for me to increase classes in some areas which makes goals somewhat unattainable. I increased qualifying scores by one third and didn't meet goals.
- Communication with NMSI (besides Sarah) has been inconsistent and frustrating.
- The liaison for the program was not accessible, did not provide information, and was basically non-existent for the majority of the school year.
- This is my third year and my experiences with CRP and expectations have always been well aligned and supported
- Being the 3rd year, I knew what to expect and that was delivered.
- I still don't feel we are doing enough to target new students to AP course. I need guidance with my teachers.
- The program has helped my students achieve academically by providing financial support.
- Our involvement required us to need thoughtful about student registration and performance in these courses. We need to provide more access to all students. I think NMSI sets the goals but doesn't provide clear rationale when goals are not attainable due to enrollment. There could be more specific recruitment strategies shared with schools.
- My expectations of the support my school would receive matched the expectations I had for the program.
- We haven't seen as big of an increase as hoped with AP exam scores, but the program has helped to improve student performance.
- It went much like the previous year did, so it matched my expectations for this year (second year). However, I don't feel the program has matched my expectations for NMSI compared to when I was involved in writing the grant.
- After working with NMSI last year, I had a good idea as to what to expect from this year's programming.
- I had spoken to the school administrator who had previously served in this capacity, and had a good idea of what to expect
- This year, the program liaison is very unresponsive and the defined goals where unrealistic for the student population that the school serves.
- Nmsi was an amazing resource for our school

- NMSI's College Readiness Program matched my expectations this year.
- My expectations were high and the program was very informative .
- The program is great, just hard to be fully implement because we are an IB School
- There are some minor issues, but overall we continue to make great strides as a district thanks to CRP
- As this is not my first year doing the grant, we knew what to expect. We were well supported through communication and partner access.
- This is my third year working with NMSI, so I already knew what to expect.

## Appendix H: 2019 PSD Survey Comments on CRP Impact on School

### More Details on Participants Rating of CRP's Impact

- I believe we are given resources to make the program successful. I think their should be a wider variety of resources that teacher can use instead of a "list".
- The study sessions and materials being able to be ordered were a big help.
- Our district is having some financial difficulties. As we move out of this difficulty we will be able to increase numbers more readily.
- I haven't thought much about what those supports, resources and development are but I believe to move our AP programs forward we will need to increase in all 3 phases.
- There are many to share. First, our Program Manager is accessible, supportive and responds to questions and inquiries quickly. Second, monthly collaboration with our AP teachers and I have allowed a strong bond to develop around a shared philosophy and goals. Finally, our students are more willing to take the leap into AP because of the support and training teachers have received
- The Saturday sessions are extremely valuable as is the help with teachers and their mock exams.
- Our AP enrollments are up, but I have a lot of students taking multiple classes instead of new students enrolling in classes.
- The trainings and collaboration with local AP teachers and administrators was helpful.
- Our students appreciated the stipends for the assessment. Our teachers have received a great deal of professional development and have collaborated with great teachers around the nation. This has dramatically changed instruction in the classroom.
- NMSI provides mentors and strong support systems for my teachers, access to supplies we normally wouldn't have had funding for, and professional development for my staff.
- I do think that most students can be successful on an AP exam at this school, but a few are too far behind academically to catch up sufficiently.
- The NMSI money encouraged us to launch AP CSP and Computer Science A sooner than we likely would have done so, so that is why I agree it has increased our student access. The non-monetary sources of support have not really lived up to expectations for increasing student performance. I don't feel like I have access to a network beyond my school district.
- Having more access to additional resources has been fantastic. I really appreciate the additional training for teachers and students as well as the wealth of resources.
- I have gotten good feedback from teachers about the NMSI training. While they appreciate the funding for additional supplies and resources, the process is somewhat cumbersome.
- The budget is not realistic for the number of students enrolled in the courses. The enrollment efforts were naturally being performed at the school prior to CRP and

NMSI. When we try to discuss critical issues and how the cookie-cutter models won't work at a school with traditionally high enrollment in AP courses, it feels as though the organization believes we are just not implementing their models with fidelity.

- Absolutely. Support is always a text or email away and Will from nmsi has been amazinf
- All statements that I checked as "strongly agree" are 100% true.
- All have improved dramatically thanks to CRP
- Perhaps the best aspect for our teachers, was that it provided them opportunities to work with other teachers that do the same thing they do.
- NMSI pushes us to push ourselves and so our school takes advantage of the resources that NMSI provides for us and makes us aware of.

## Appendix I: 2019 PSD Survey Factors that Affected CRP Implementation

**Contributed to Success**

- Support and training was a great help. Having the ability to get multiple teachers trained have educated them on how AP should be implemented.
- The teacher training coupled with each component and support of the program manager all had a significant impact in our success. It was difficult to choose the LEAST effective component because they are all important.
- SSS, Summer Institute an Mock exam feedback
- Definitely appreciate the financial incentives and resources (stipends, materials, money for supplies)
- Teachers enjoyed the professional development experiences, content, and strategies.
- The AP teacher workshops significantly helped.

**Impeded Implementation**

- Communication. It is inconsistent and confusing.
- Small enrollment in our building required us to pull from same small group of students--the same group of kids took all of our AP courses.
- Teachers need continued PD on recruiting students. Many still think AP students must meet a certain profile.
- The elimination of supports after three years is challenging. I am unsure why schools can't continue to receive a discount on student tests.
- District budget cuts
- We have been impeded by competing dual-enrollment programs, as well as a decrease in the reading and math level of incoming 9th graders. Also, teacher transitions in some AP courses caused slight difficulties.
- I feel like my biggest disappointment with CRP is that it hasn't seemed to have had an impact on teachers who had previously been teaching the AP course. I wonder if it would have been more helpful if NMSI had a menu of options for teachers, and allowed teachers to customize their interventions, so that they felt like they had buy-in for the program, as well as choices. I think some of the teachers feel like the program is just another box to check, and so haven't done more than the bare minimum. I hope if they'd been given choices rather than being told what to do, they would have embraced the interventions more. Also, the mentor was not helpful in most cases. It seems like the mentors have large case loads, and not a lot of time to give the teachers. I know some of my teachers would have preferred, and likely benefitted from a mentoring session that occurred a few times a month, rather than a summer training and two weekends a year.
- Some of the data requests from NMSI are not aligned with readily available formats;
- Being a IB School impeded on successfully implementing the CRP

## Appendix J: 2019 SC Survey Suggestions for Improving SSS Attendance

**Recommended Methods for Increasing SSS Attendance**

- The only pitfall is the month where sessions are every Saturday. That is taxing for students who participate in multiple sessions.
- NMSI requires the dates to be scheduled so far in advance, other things that are more important to students (like prom) haven't been scheduled yet. So it is hard to plan around those dates for the spring study sessions.
- More consistent quality with presenters, more engaging review lessons, and clearer outcomes for each session.
- coordinators should work with other teachers to prevent double booking students with conflicting events.
- If students attend all 3, I personally provided tickets to Cedar Point Amusement Park
- Don't do it on a Saturday.
- Preview of topics prior to the event
- I suggest teachers offer incentives for participation.
- I am going to take a more active role at the beginning of the school year to coordinating with the teachers of record. HAIS will offer incentives. Need marketing material to post around school, in classes, and email parents.
- day and length
- Provide recruitment best practices
- Attendance from my school is pretty good. I wouldn't change anything.
- Short of paying students a small stipend for attending, nothing. Students must take responsibiltiy for their own education.
- More incentives and looking at different dates in advance. I believe the last SSS was on the same day as the ACT exam.
- To have it during after school activities instead of Saturday mornings.
- Tying a grade to attendance with an alternative assignment for those with pre-approved valid excuses. The students will not take it seriously if the teachers don't.
- Shorter sessions (3 hour sessions instead of 4 hour sessions)
- Make the study sessions closer to the test. There is no urgency for students to attend a session in the fall.
- We provide comp days to students who attend, meaning they can miss up to 2 class periods in a given class after the AP exam for every study session they attend in that subject area.
- Have each school host a day.

- I am not sure, working on that. Our students are burnt out by the time Saturday comes from working hard during the week. Also several students are involved in other things on Saturdays
- no incentives for passing without meeting the 3 minimum
- Flexible dates if students are in activities. We tried to work around activities but i still think there were conflicts.
- Bonus Points seem to have the greatest influence on student attendance
- Potentially a financial incentive for attending. Also ensuring teachers are making the push to encourage students to attend.
- Differentiated sessions may be appropriate. The first session is very basic and our AP teacher had already covered the material, and this may have discouraged the students from future sessions.
- It should be 3 hours max
- Plan study sessions on non act days
- Students don't attend because of other school conflicts, which can't always be avoided.
- normalizing food/beverage offered to students so they know whether or not they can expect lunch afterward
- increase the reimbursement of $3 per student for food. maybe provide lunch afterwards
- Teachers offer incentives, but sometimes the students have other commitments (sports, musical, etc.)
- As the year has progressed there has been better teacher buy in. Teacher buy in is key.

## Appendix K: 2019 SC Survey Factors that Hindered or Assisted CRP Implementation

**What helped or impeded ability to fulfill role:**

- Communication from NMSI staff (aside from Sarah) has been inconsistent and confusing.
- We need van size transportation. We are a small school and we will never fill a regular size bus.
- Communication reminders helped.
- I also stated these reasons in the Partner School Director survey...competing dual-enrollment programs, a decrease in the reading and math levels of incoming 9th graders, and some teacher turnover in AP classes all make the CRP more challenging for our school to effectively implement.
- Ashley Pita was extremely helpful and responsive when organizing and preparing for SSS.
- I'm good as I am very committed to the College Readiness Program.
- Muriel Alim was a factor that significantly helped my ability to fulfull the role. Daily school operations made it difficult to fulfill the role.
- Some of us have other duties on campus like GT/Magnet Coordinators and we have events on the same day. If we don't attend the Saturday sessions we don't get paid. We do NMSI duties on campus that are not taken into consideration when the stipend is issued.
- Buying food in bulk is the worst!
- Overall it is just a capacity issue. I believe and value our partnership with Nmsi...it is a lot with all other expectations and responsibilities on my plate
- Communication was a bit spotty this year, which led to some last minute signing up and/or changes.
- The math teachers and school coordinators came late or didn’t come at all.
- I think it's a pretty good system. I have received materials much earlier this year, which is okay for sorting but I've had to run copies because not enough materials were sent.
- Sometimes more organization and a quicker response to emails would have been helpful.