

# UC Riverside

## UC Riverside Previously Published Works

### Title

Life on the Frontier of AP Expansion: Can Schools in Less-Resourced Communities Successfully Implement Advanced Placement Science Courses?

### Permalink

<https://escholarship.org/uc/item/7k79s7t8>

### Journal

Educational Researcher, 48(6)

### ISSN

0013-189X

### Authors

Long, Mark

Conger, Dylan

McGhee, Raymond

### Publication Date

2019-08-01

### DOI

10.3102/0013189x19859593

Peer reviewed



# HHS Public Access

Author manuscript

*Educ Res.* Author manuscript; available in PMC 2020 June 03.

Published in final edited form as:

*Educ Res.* 2019 August ; 48(6): 356–368. doi:10.3102/0013189x19859593.

## Life on the Frontier of AP Expansion: Can Schools in Less-Resourced Communities Successfully Implement Advanced Placement Science Courses?

Mark C. Long<sup>1</sup>, Dylan Conger<sup>2</sup>, Raymond McGhee Jr.<sup>3</sup>

<sup>1</sup>University of Washington, Seattle, WA

<sup>2</sup>The George Washington University, District of Columbia

<sup>3</sup>Equal Measure, Philadelphia, PA

### Abstract

The Advanced Placement (AP) program has undergone two major reforms in recent decades: the first aimed at increasing access and the second at increasing relevance. Both initiatives are partially designed to increase the number of high school students from low-income backgrounds who have access to college-level coursework. Yet critics argue that schools in less-resourced communities are unable to implement AP at the level expected by its founders. We offer the first model of the components inherent in a well-implemented AP science course and the first evaluation of AP implementation with a focus on public schools newly offering the inquiry-based version of AP Biology and Chemistry courses. We find that these frontier schools were able to implement most, but not all, of the key components of an AP science course.

### Keywords

Advanced Placement; curriculum; experimental research; high schools; implementation; mixed methods; multisite studies; postsecondary education; program evaluation; science education; teacher context

---

Advanced Placement (AP) courses have become a desired amenity for many high school students with intentions of someday earning a bachelor's degree. The College Board (the "Board" for brevity), the organization that administers the AP program, seeks to make the courses accessible to all students and relevant to their college and labor market experiences. In fact, accessibility and relevance have been at the forefront of AP programming for the last several decades. The expansion of AP course availability, and the introduction of a new framework for AP science, has raised questions about whether teachers in less-resourced communities can successfully implement the AP program.

At the AP program's inception in the mid-1950s, the courses were available primarily to students with high achievement in wealthy schools (General Education in School and College, 1952; Lichten, 2010). The Board has since implemented a number of initiatives aimed at removing disparities in access to AP courses. One major change has been the

---

Board's explicit *AP for All* message that clarifies for students and educators that AP is not just for students who always score at the top of their class, but for all students who are willing to accept the challenge of a college-level course (College Board, 2019b). The Board soon followed with the *All In* initiative, which encourages schools to test all students with the Preliminary Scholastic Aptitude Test (PSAT) so that they can identify students with the potential to succeed in an AP class (College Board, 2019a). These efforts have been backed by district, state, and federal policymakers with subsidies to low-income students to take AP exams and, in some localities, requirements that schools offer AP classes in order to meet accountability benchmarks, which vary across states (Adelman, 2006; Education Commission of the States [ECS], 2016; Holstead, Spradlin, McGillivray, & Burroughs, 2010). The George W. Bush administration gave a substantial boost to the AP program, particularly its math and science courses, through the American Competitiveness Initiative. As part of a larger effort to strengthen science, technology, engineering, and mathematics (STEM) education, the initiative poured \$122 million into training new AP math and science teachers and subsidizing AP exams for low-income students (The White House, 2006).

These initiatives, along with an increase in the use of AP in postsecondary admissions decisions, have led to a rapid expansion of the AP program (Judson, 2017; Judson & Hobson, 2015). Since 1990, AP courses experienced an annual growth rate of 8.5% and are now found in over 70% of U.S. public high schools (Malkus, 2016). Because nearly every large U.S. secondary school offers AP, 90% of all high schoolers attend a school with at least one AP course (Malkus, 2016). The race and income gaps in access to AP that are driven by differences in school course offerings are almost a problem of the past. In science, for example, the share of White, Black, and Hispanic students attending schools that offer an AP course are 73%, 72%, and 78%, respectively (Malkus, 2016).<sup>1</sup>

In addition to increasing the accessibility of AP, the Board has restructured many of the courses to increase their relevance to students' college and labor market experiences (College Board, 2017). Today's AP classes are supposed to place less emphasis on memorization and greater emphasis on "discipline-specific inquiry, reasoning, and communication skills," akin to the kind of learning that occurs in a classroom at a selective university (College Board, 2017).

These changes have been most pronounced in AP science. In collaboration with the National Science Foundation (NSF), the National Research Council (NRC), and educators across the nation, the Board revised its AP science curriculum to make it more consistent with a conceptual framework laid out by the NRC in 2012 (NRC, 2012). The new AP science courses focus on exposing students to the real-world practice of science. In an AP science classroom, students are supposed to direct the inquiry; engage in small group experimentation; and be guided by teachers to ask questions, develop hypotheses, design experiments, and explain their observations to one another (College Board, 2011a, 2011b). AP instructors are also encouraged to learn different modes of instruction in order to

---

<sup>1</sup>Advanced Placement (AP) course-taking gaps by race and income and other demographic indicators persist within schools and are driven by a number of factors. See, for example, Klopfenstein (2004); Conger, Long, and Iatarola (2009); Rodriguez and McGuire (2019).

effectively teach students with different learning styles. This includes greater use of technology to help students explore relationships in the data and communicate their findings. Science educators posit that this inquiry-based approach will better prepare students for careers in STEM fields and reduce barriers to STEM engagement often experienced by students from low-income backgrounds and underrepresented minority groups (Kurth, Anderson, & Palincsar, 2002; Litzler, Samuelson, & Lorah, 2014).

Some critics argue that AP growth has led to a watering down of the curriculum or, at the very least, a level of quality that is completely unknown (Bowie, 2013; Farkas & Duffett, 2009; Lichten, 2010; Tai, 2008; Tierney, 2012). Regarding the expansion that occurred through the American Competitiveness Initiative, for instance, Tai (2008) writes, “It is very likely that, rather than teaching the high-level science and mathematics required to pass the AP exams, the new teachers hired through the American Competitiveness Initiative would find themselves focusing on remediation.” (p. 4). Regarding the new inquiry-based curriculum, others worry that schools with limited capacity will have a particularly difficult time moving away from rote memorization and lecture-based approaches to instruction (Schneider, 2013). Much of the concerns stem from AP teachers themselves, who report that many of the students enrolling in AP classes are underprepared and that as the AP program strives for equity in access, it may compromise its excellence in delivery (Farkas & Duffett, 2009).

To our knowledge, these concerns have never been rigorously explored. There is a modest body of research on the *impact* of AP courses and exams on student outcomes (e.g., Avery, Gurantz, Hurwitz, & Smith, 2017; Jackson, 2010, 2014; Klopfenstein & Thomas, 2009; Smith, Hurwitz, & Avery, 2017; Warne, Sonnert, & Sadler, 2019). Yet we could find no research on the fidelity with which AP courses are implemented. This article offers the first such evaluation of AP implementation with a focus on public schools offering the inquiry-based version of AP Biology and Chemistry courses for the first time. Our research is part of a larger experimental study that we launched in 2012 aimed also at evaluating the impact of taking an AP science courses on student outcomes. This larger study will provide the first credible estimates on the effect of AP science course taking on important student outcomes, including science skill, STEM interest, and college enrollment.<sup>2</sup> Here, we provide a theory that describes the ideal context (e.g., well-prepared and motivated teacher) and treatment components (e.g., college-level, inquiry-based, syllabus) for a well-implemented AP science course. We then evaluate participating schools on the extent to which they were able to reach these ideals with fidelity.

---

<sup>2</sup>Estimation of main impacts on some of these some of these outcomes, and heterogeneity in impacts, are included in our companion manuscript: Conger, Kennedy, Long, and McGhee (in press). “We find suggestive evidence that taking an AP science course increases students’ science skill and their interest in pursuing a science, technology, engineering, and mathematics (STEM) major in college. AP course-takers also have lower confidence in their ability to succeed in college science, higher levels of stress, and worse grades than their control counterparts” (abstract). All of the prior research on AP impacts is observational, and most is also correlational. Of those studies that aim to produce more causal estimates by controlling for confounding variables or relying on variation in exposure to AP courses, the findings are mixed; some studies find positive effects of access to AP on outcomes such as college enrollment (e.g., Jackson, 2010, 2014), whereas other studies show no AP impact on outcomes such as STEM interest (e.g., Warne, Sonnert, & Sadler, 2019).

## Theory of AP Science Courses

### The AP Label

The process by which a course is designated “AP” involves two steps. Teachers who plan to offer an AP course are encouraged (though not required) to attend a professional development training. The Board and other independent agencies offer several workshops, with the most extensive training being the AP summer institute, a week-long training that is led by an experienced AP instructor. Teachers are then expected to develop their syllabi for the course and submit them to the Board for review. A team of auditors at the Board reviews each syllabus and grants permission to a school to label the course as AP on course catalogs and student transcripts once the syllabus has been approved. The Board provides several resources to help teachers prepare their syllabi, including the curricular and resource requirements for syllabus approval (e.g., use of a recently published textbook).<sup>3</sup>

### Inputs for a Well-Implemented AP Science Class

Our evaluation is guided by a theoretical framework that describes the core components required for successful implementation of an inquiry-based AP science course. We developed our framework by reviewing NRC recommendations for science education (e.g., NRC, 2002, 2012) and documents published by the Board regarding the curricula (e.g., College Board, 2011a, 2011b), and in consultation with the curriculum development team at the Board.<sup>4</sup>

Figure 1 provides a list of components that the Board deems important to support teaching and learning in an AP science course (“context”) and the core components expected of a well-implemented course (“treatment”). The context column lists the ideal set of conditions that exist outside of the classroom and that support effective instruction, including (a) well-prepared and motivated teachers, (b) well-prepared and motivated students, (c) sufficient planning time and resources, and (d) supportive school and district conditions.<sup>5</sup>

With these key contexts in place, a well-implemented course should result in the following five treatment components: (a) a college-level, inquiry-based, syllabus; (b) coverage of the major science practices and learning objectives; (c) an academically challenging curriculum; (d) project-based and independent classroom activities; and (e) integrated use of technology. Upon successful course implementation, the Board expects students to obtain a high level of competency and interest in the subject matter and thus be prepared for the rigors of college-level work.

This theory rests on the assumption that the existing courses that students would otherwise take do not offer the level of rigor, inquiry, or direct connection to postsecondary education that the AP course offers. Depending upon school offerings, students who seek demanding

<sup>3</sup>For more detail on the audit process, see <http://www.collegeboard.com/html/apcourseaudit>. This audit process leaves a lot of discretion to schools to shape their curriculum. For example, see <https://apcentral.collegeboard.org/pdf/ap-biology-syllabus-development-guide.pdf?course=ap-biology>.

<sup>4</sup>LaTanya Sharpe, the College Board’s Associate Director of AP Science, was our primary resource on the AP science curriculum.

<sup>5</sup>Individual districts and teachers may have different views about the resources and materials necessary to successfully implement an AP science course. The resource requirements listed by the College Board are flexible and are listed on p. 123 of College Board (2015) for AP Biology and p. 112 of College Board (2014) for AP Chemistry.

instruction have three other options. Most high schools offer honors courses, which are intended to provide more rigor than a regular high school course, but not necessarily at a college level. Some students can also enroll in dual enrollment or dual credit courses, which are taught by college instructors often at a nearby college or online. In the most recent national survey, high schools reported approximately 2 million enrollments in dual credit courses (Thomas, Marken, Gray, & Lewis, 2013). Finally, some schools have the International Baccalaureate Diploma Programme (IB DP), which was originally designed for international schools and aimed at teaching students critical thinking skills and knowledge of world affairs. The IB DP, which includes robust authorization policies, remains relatively uncommon in the United States with less than 5% of high schools offering IB DP in 2016 (International Baccalaureate Organization, 2016).

## Study Overview

### Evaluation Design and School Recruitment

We designed the study as a randomized controlled trial in order to obtain credible estimates of treatment effects on student outcomes. To achieve randomization, we recruited schools from across the nation that had not offered AP Biology or AP Chemistry in recent years; were willing to add such a course and comply with study protocol; and had more eligible students than could be served in one class, where criteria for being “eligible” was defined by each high school. Twenty-three schools participated with 12 schools adding AP Chemistry, 10 schools adding AP Biology, and 1 school adding both courses.

For at least 1 year (and no more than 3 years), each participating school provided us with a list of students that the school deemed eligible to take the new AP course. Upon receipt of signed guardian consent and student assent forms, we then randomly assigned a subset of participating students to the offer of enrollment in the new course. Students who were not offered enrollment were permitted to enroll in any other courses offered in the school. The study includes two waves of participating schools (those who joined in 2012–2013, and those who joined in 2013–2014) and 1,820 students. The unit of analysis for this implementation evaluation is the 24 unique school courses (23 schools offering either biology or chemistry and 1 school offering both courses).

### Implementation Evaluation Data Sources

We evaluated implementation quality with five quantitative and qualitative sources of data. First, we surveyed teachers offering the new AP courses online at the conclusion of the course with questions about their educational background, professional experiences, and professional development; past and present instructional practices generally and around science specifically; participation in the required AP training; ability to cover the content of the AP course; and coaching, mentoring, and other professional community supports. The response rate for this survey was 100%. In the analyses below, we collapse the teacher-level data to the school-course level by computing the mean response, yielding 24 observations. Second, we obtained course syllabi from teachers and scored each syllabus according to coverage of the major science practices and learning objectives expected of an AP course. Third, we conducted 61 interviews of teachers, principals, and other school and district staff

focusing on the challenges they faced in fielding the new AP course, including any resources they required above what they expected. Our interviews also touched upon district policies, school culture, administrative support, and coordination mechanisms.

In addition to these three course and teacher levels of data, we collected two sources of data on students. We administered paper and pen surveys to participating students in the spring of each year, prior to the Board's administration of AP exams. This survey had a response rate of 78% (1,417 students). The survey included implementation-relevant questions related to students' perceptions of the AP science course (or their most recent science course for control group students), including the degree of rigor; use of technology; and use of independent, project-based activities. We also obtained student transcripts to determine whether participating students had completed the recommended prerequisites for the AP course.<sup>6</sup>

### Characteristics of Participating Districts, Schools, and Teachers

Figure 2 shows the geographic distribution of the 11 participating school districts, which are primarily concentrated in the western, southern, and eastern regions of the country.<sup>7</sup> Relative to districts across the nation, those participating in the study tend to be in neighborhoods with middle or lower levels of socioeconomic status and to educate students who score below average on tests in earlier grades, as illustrated in Figure 3.

Participating schools are larger and more likely to educate students who are eligible for free or reduced-price lunch, Black, and Hispanic than other schools (Panel A of Table 1).<sup>8</sup> These are exactly the type of new frontier schools that many worry will be unable to implement an AP science course with a high level of integrity. By "frontier," we mean the types of schools who may be next to expand into AP science course offerings. Prior research on advanced course offerings suggests that offerings are higher in larger schools with fewer students who are eligible for subsidized meals (Iatarola, Conger, & Long, 2011; Monk & Haller, 1993). While our participating school districts have *higher* expenditures per pupil, this is commonly observed in "less-resourced" urban districts where higher costs of living lead to higher teacher salaries, more students requiring subsidies for free and reduced-price lunch, and thus higher expenditures. Thus, when we use the term "less resourced," we mean districts that are in areas with middle or lower socioeconomic status and with higher shares of students who are receiving free or reduced-price lunch.

Reflecting the school demographics, participating teachers are slightly younger; less experienced; and more likely to be female, Black, Asian American, and Hispanic than high school science teachers nationally (Panel B of Table 1). Study teachers are more likely to

<sup>6</sup>The College Board recommends Chemistry I and Algebra II as prerequisites for AP Chemistry and Biology I and Chemistry I for AP Biology, with no additional requirements beyond these prerequisites. Our analysis of transcripts evaluated whether the students completed these courses. The participating schools, however, may have their own standards for eligibility, including a different set of prerequisite courses or no prerequisite courses.

<sup>7</sup>Participating districts include Anaheim Union High School District, California; East Side Union High School District, California; Lynwood Unified School District, California; Jefferson Parish, Louisiana; Education Achievement Authority, Michigan; Charlotte-Mecklenburg Schools, North Carolina; Winston-Salem/Forsyth Schools, North Carolina; Cranston Public Schools, Rhode Island; El Paso Independent School District, Texas; Metropolitan Nashville Public Schools, Tennessee; and Richmond Public Schools, Virginia.

<sup>8</sup>Source: 2013–14 Common Core Data, <https://nces.ed.gov/ccd>; EDFacts, <https://www2.ed.gov/about/inits/ed/edfacts/index.html>.

hold an undergraduate major in a STEM field than other high school science teachers, yet far less likely to hold a masters' degree and slightly less likely to have earned a teaching credential in science. Most of the participating teachers had previously taught a higher-level course (mostly honors), yet only 47 percent of them had previously taught an AP course. Our evaluation consequently generalizes to a population of teachers who are relatively new to the AP science curriculum and who have generally not received graduate training in STEM.

## Methods for Evaluating Implementation

Our methods seek to answer the following two questions:

1. To what extent are the context and treatment components shown in Figure 1 present in these frontier schools?
2. Does taking AP science, rather than other science courses, cause the student to receive a significantly and substantively different treatment?

To answer the first question, we construct indices for each of the four context and five treatment components listed in Figure 1. We take two steps to compute these indices. First, we convert all ordered categorical variables into 0–1 ranges by assuming that the lowest possible response equals 0, the highest possible response equals 1, and intermediate responses can be linearly mapped. For example, one of the indicators we use to evaluate teacher preparation comes from the survey question “typical instruction involves student self-directed learning,” which contains four possible responses ranging from *not at all* to *a great deal*. We recode these responses to 0, .33, .67, and 1. We then compute the simple average of the individual indicators for each element of context and treatment. Appendix Tables A1 and A2 list all indicators used to measure each component (where indicators were drawn from multiple data sources, including course syllabi, teacher surveys, and student transcripts). For instance, the index to measure “well-prepared and motivated teachers” is the average of the 16 indicators listed in Appendix Table A1. We recognize the strong assumptions that we are making (e.g., that each indicator has equal worth, and that ordered categorical variables can be represented linearly). Thus, in the discussion below, we discuss results of individual indicators where appropriate.

To answer the second question, we use the results from the survey of participating students and estimate the effect of taking the AP science course on students' perceptions of course content and rigor (also referred to below as treatment-control contrast). We estimate this contrast with a standard instrumental variable specification, where the randomized offer of enrollment in AP is used as an instrument for actual enrollment. Using an instrumental variable is needed as enrollment in the class was not mandated for those who were randomly assigned to the treatment group. Rather, students who were offered enrollment then chose to accept or reject that offer. Additionally, a small group of students were allowed by school administrators to take the AP course despite being randomly assigned to the control group. This imperfect compliance to randomization means that any differences in the experiences of students who took the class versus those who did not might capture a combination of the



true differences (e.g., the AP science course really is more intellectually challenging) and differences in the types of students in the two classes.<sup>9</sup>

To overcome this problem, we use two-stage least squares to estimate the model for all outcomes. The local average treatment effect (LATE) estimate (the causal effect of the course on those who complied with their random assignment) is given by  $\beta$  in the first equation that follows (Angrist, Imbens, & Rubin, 1996; Bloom et al., 1997; Imbens & Angrist, 1994).<sup>10</sup> As noted by Angrist and Pischke (2009), “We can think of instrumental variables as initiating a causal chain where the instrument [*Offered<sub>i</sub>*] affects the variable of interest, [*AP<sub>ij</sub>*], which in turn affects outcomes, [*Y<sub>ij</sub>*]” (p. 151):

$$Y_{ij} = \alpha_j + \widehat{AP}_{ij}\beta + \mathbf{X}_i\gamma + \epsilon_{ij}, \quad (1)$$

$$AP_{ij} = \delta_j + Offered_i\theta + \mathbf{X}_i\mu + \epsilon_{ij}, \quad (2)$$

$Y_{ij}$  is the treatment indicator response of student  $i$  in school  $\times$  cohort stratum  $j$ .  $\widehat{AP}_{ij}$  is the fitted value based on the estimates of the parameters in Equation (2).  $\mathbf{X}_i$  is a vector of pretreatment covariates, including age, math and reading exam scores from 8th and 10th grade (standardized and averaged for math and reading separately); cumulative grade point average prior to the year when the AP science course was offered; and indicator variables for female, racial group (Asian American; Black; or Hispanic, Native American, or Multiracial), disability, gifted, English Language Learner, eligible for free or reduced-price lunch, home language is not English, took recommended prerequisite courses.<sup>11</sup>  $AP_{ij} = 1$  if the student enrolled in the AP science course.  $Offered_i = 1$  if the student is randomized into the treatment group.<sup>12</sup>  $\alpha_j$  and  $\delta_j$  school by cohort fixed effects. For statistical inference, standard errors are clustered by school  $\times$  cohort.<sup>13,14</sup>

Given the possibility of nonrandom attrition due to some student participants not completing the survey, we weight all regressions by the inverse of the probability of completing the survey conditional on student characteristics (unweighted results, available from the authors, are quantitatively similar).

<sup>9</sup>For instance, suppose the control group students who managed to get into the AP science course are more likely to report that the course is challenging because of some underlying attribute about them (motivation, interest, ability) that renders them different from the control group students who complied with their random assignment status.

<sup>10</sup>Similar approaches are used in recent studies, including Abdulkadiro, Lu, Pathak, and Walters (2018); Boatman and Long (2018); Hull, Hinerman, Ferguson, Chen, and Näslund-Hadley (2018); and Xu, Solanki, McPartlan, and Sato (2018).

<sup>11</sup>Controlling for these variables improves precision in the estimates. On the majority of these precourse characteristics, there was good balance between those that randomized the offer of enrollment and those that did not (meaning, the two groups were statistically equivalent). One notable exception was on students' reading exam scores, where treatment group students scored 0.09 standard deviations higher ( $p$ -value = .02). Thus, controlling for this variable is particularly important as it adjusts for the chance imbalance.

<sup>12</sup>If we were to replace  $\widehat{AP}_{ij}$  in Equation (1) with  $Offered_i$ , the resulting coefficient,  $\beta$ , would reveal the “Intent to Treat” (ITT) effect—that is, the effect of our intention to treat the student by enrolling the student in the AP course.

<sup>13</sup>The difference in AP science course-taking rates between treatment and control groups is captured by  $\theta$  in Equation (2). Fifty-eight percent of the students who received an offer chose to enroll, and 17% of the control students were allowed by school administrators to enroll in the course; that is, the study's randomized offer of enrollment boosted the likelihood that the student enrolled in the AP science course by 41 percentage points (i.e.,  $\theta = 41$ ,  $SE = .06$ ). This first-stage instrument is very strong, with an  $F$ -statistic of 46, which is far above the commonly cited rule-of-thumb threshold of 10 (Staiger & Stock, 1997).

<sup>14</sup>Relative to control group “compliers,” members of the control group who did not comply with their assignment to the control group were less likely to be White.

Finally, to answer both questions above, we supplement our quantitative metrics with qualitative insights from our teacher and principal interviews.

## Results

### Evaluation of Context

Figure 4 provides the distribution on the index scores used to measure the four key context components recommended for successful AP course implementation. Each marker represents one school-course observation, and the gray bars show the interquartile range of these observations.

We use a number of metrics to capture the degree of teacher preparedness and motivation, including their education and professional development; the degree to which their typical instruction resembles the AP inquiry-based approach; and whether they attended an AP summer institute training along with their evaluation of the training. The mean for this element of context was .69 (i.e., the mean was more than two-thirds of its maximum value). Though most teachers ranked relatively high in preparation, the indicators that contributed most to a lower ranking on this component included whether the teacher held a master's degree, whether their typical instruction involved class discussions and debates, and the overall similarity between their typical instructional approach and the AP science approach. For instance, on a scale ranging from 1 (*not at all similar*) to 3 (*very similar*), teachers responded with an average of 2 to the question of the similarity between their typical instructional approach and the AP approach.

In our interviews with teachers, we recorded a very high level of motivation for teaching the AP courses. Some teachers specifically sought out the opportunity to teach the new AP course because of a desire to interact with students at a higher level and for students to be successful in college and career with one teacher commenting about the course: "it's challenged my mind, its difficult material and I haven't done it in a long time." Consistent with these interview reports, most teachers (87%) reported in the survey that they would be willing to teach the course again if given the opportunity.

To capture the degree of preparation and motivation of the students, we measured the percent of students who completed the College Board's recommended course prerequisites, their average GPA in those prerequisites, and the teachers' perspective on the extent to which students' incoming knowledge and ability helped or hindered course implementation. Most of the students had completed the prerequisite courses (88% in science and 97% in math) and received an average GPA of 3.25 in those courses. And the average across schools on this measure of context was relatively high (a mean of 0.72).

Despite having met the prerequisites, when we asked teachers the extent to which students incoming knowledge hindered or helped with course implementation (on a scale of 1 = *significant hindrance* to 5 = *significant help*), the mean response was a 3. In interviews with teachers, we learned that there were two major factors leading some of them to give lower scores to student preparedness. First, some teachers explained that the variation among students in the classroom hindered effective course instruction, with one teacher reporting,

“The valedictorian is in my class and then there’s some that have a C average.” This teacher explained that she wanted “to move at a faster pace than [the students] were willing and ready for.” Second, several teachers and school leaders felt that many students were not well prepared for the inquiry-based nature of the course. At one school, for example, the principal noted that the school does not offer an honors biology class to help prepare students for the skills required of AP. He explains,

Most of the students in AP Bio are co-enrolled in Bio for the first time. They’re really coming in at a deficit in terms of prior knowledge. Even those who are seniors and took it as freshman, it’s pretty clear that whatever experience they had as freshman didn’t give them the prerequisite skills that they need to be successful.... The way that’s played out is we’re spending more time than I’d have liked to on developing the content knowledge that they need in order to really be asking meaningful questions in an inquiry-based laboratory.

Though teachers and school leaders discussed the challenge of implementing the course when some students lack the content knowledge, they mainly described the challenge of students’ lacking exposure to higher-order thinking and inquiry skills. As one department chair noted: It is a huge shift to go from “spitting out knowledge” in previous courses, to the AP course being about “what you can show me.” Another respondent commented that students had spent years getting used to “cookbook”-style labs, in which directions are explained step-by-step and no inquiry is involved.

Across these schools, there was a wide range of responses to indicators of “sufficient planning time and resources,” with a mean indicator value of .64. These schools experienced more challenge with “planning time” (indicator = .49) than with and “resources and materials” (indicator = .76).<sup>15</sup> And when asked about the level of support from school/district leaders in “curriculum development,” “instructional materials,” and “pedagogy,” the average responses were each “some, but not enough, support” (indicators = .24, .31, and .25, respectively).

The largest context deficiency was in school and district conditions. The vast majority of AP science teachers reported low levels of support from the school and district, with a mean indicator value of .38. Notably, only about one-fourth of teachers in these courses reported that “inquiry-based learning” and “science content or facts” were regularly discussed during team meetings or reported having “a mentor or coach at school that provides support.” These reports align with surveys of teachers nationally where those in large schools with higher shares of students eligible for free or reduced-price lunch tend to report lower levels of support and encouragement from their administrators than teachers in other schools (U.S. Department of Education, n.d.).

### **Evaluation of Treatment: Teacher Survey and Course Syllabi**

Given this contextual mix of reasonably well-prepared and motivated teachers and students, combined with varying levels of planning time and resources, and low reported levels of

---

<sup>15</sup>We did not specify a list of resources in our survey questions as the College Board’s guidance on required resources gives schools flexibility to tailor their course.

support from the school and district, it is uncertain whether the treatment can be adequately implemented. Figure 5 presents evidence on the distribution of treatment fidelity based on our review of course syllabi and teacher survey responses. We find mixed results.

Of the 24 courses, 23 had their syllabi audited by the Board. Yet we find a wide range of success with integrating key concepts. Most notably, we find that, for the typical syllabus, there was only isolated integration of practices and communications with content and integration of issues around society, technology, and innovation with content, thus yielding fairly low average indicators (.46 and .44, respectively).

On the plus side, teachers reported high levels of success in covering major science practices and learning objectives and delivering an academically challenging curriculum. More than half of the teachers reported that their AP course was “a lot more challenging” than an honors course. Yet, consistent with our evaluation of syllabi, teachers reported mixed results on their ability to incorporate project-based and independent classroom activities and integrated use of technology. Teachers’ greatest success was in their ability to allow students to work in small groups, with the average response slightly above “frequently” (indicator = .79). Teachers were less likely to provide time for students to “read book (non-textbook) or magazine about science” (indicator = .36), and, on average, teachers only “occasionally” provided time for students to “present what they learned about a topic to the class” (indicator = .48). Moreover, on average, teachers were only “occasionally” able to integrate technology to “practice concepts,” “conduct interactive simulations,” “create graphical presentations,” and “develop collaborative projects and/or group presentations” (average indicator = .55).

Thus, whereas AP science teachers in these frontier schools felt that they were able to deliver an academically challenging curriculum, and were mostly successful in covering major science practices and conveying their learning objectives, they expressed less success incorporating project-based and independent classroom activities and integrated use of technology.<sup>16</sup> This deficiency would likely lessen the exposure of these AP science students to the real-world practices of science, as was desired by the NRC, NSF, and Board.

### **Evaluation of Treatment: Treatment-Control Contrast From Student Survey**

These teacher responses are perfectly reflected in the results of our comparison between treatment and control group students in their perceptions of the rigor and content of their science courses (see Table 2). The first column of Table 2 shows the mean response of control group students who complied with their assignment to the control. The second column of Table 2 shows our estimate of the effect of taking AP science relative to other science courses offered by the school (including honors and other advanced offerings). As described in the methods section, this is referred to as the LATE, the effect of the treatment on the population of compliers.<sup>17</sup>

---

<sup>16</sup>Such challenges are likely to be common across schools in less-resourced communities. A useful extension of our work would be a study comparing the ability to incorporate technology in less-resourced schools that have AP science courses to those less-resourced schools that do not offer AP science.

<sup>17</sup>The control students who did not take AP Biology or Chemistry took a variety of alternative science courses, with the most commonly reported courses including Chemistry (13%), Physics (12%), AP Environmental Science (11%), Biology (10%), Honors Biology (9%), and Anatomy/Physiology (9%).

Students who took AP science were far more likely than control group compliers to report that their class was “intellectually challenging” (by 37 percentage points), that they “often receive homework” (by 39 percentage points), to have a teacher that “set high standards” (by 19 percentage points), and to be among classmates who were “driven to succeed” (by 24 percentage points).<sup>18</sup> Thus, both teachers and students agree that the AP science class was academically challenging.

Furthermore, AP science students were significantly and substantially more likely to “participate in hands-on learning,” “design own projects or experiments,” and “work in small groups” than their peers in non-AP science classes. However, there was no significant differences between AP and non-AP science students in their likelihood to often “work independently,” “present what [they] learned,” “apply knowledge to solve new problem[s],” or integrate technology. Thus, both teachers and students agree that there were some limitations in the ability of the AP science class to fully deliver the independent, project-based, and technology-incorporating experience that was intended by curriculum developers to generate scientific inquiry skills.

## Discussion and Conclusions

Our implementation evaluation of AP science courses in less-resourced communities reveals that many of the courses are implemented with a high level of academic and intellectual rigor compared to the other courses offered in the schools. Teachers of AP also reported a high level of motivation to teach the courses, and school leaders reported high levels of motivation for sustaining and expanding the AP courses.

At the same time, the main obstacle to effective implementation of an AP science course seemed to be in the inquiry-based nature of the curriculum. Most teachers felt that students were not well prepared for inquiry-based learning, and the AP treatment courses did not engage students more in these types of activities than the other science courses in the school. Part of the challenge may be in the lack of prior exposure to inquiry-based learning (on the part of both teachers and students) as well as insufficient planning time and support from the broader district/school around this relatively newer approach to science instruction.

Some teachers shared strategies to get around the challenge of student preparedness. In order to support the struggling students, one teacher had to scaffold the material for students. According to her, “Scaffolding is huge for them. To start where I can grab them.” Oftentimes this meant modifying the resources and materials that she finds online. At two schools, the teacher uses a modified version of a flipped classroom where the teacher assigns students a chapter and problems over winter break to avoid them falling behind. At another school, the teacher would implement something similar, where students complete assignments at home so that there is additional time during the school day for labs. One school also demonstrated a strong professional learning community in which school leaders and teachers held regular discussions and worked collaboratively on vertical alignment and articulation around inquiry

---

<sup>18</sup>Note that because we use a linear probability model for each outcome, predicted outcomes can lie outside the 0–1 interval, as is the case for the predicted response among treatment group members to the statement that the “course was intellectually challenging.”

and math in all the school's science classes. Yet these examples were rare, and teachers clearly need more training on how to foster inquiry-based learning. It would be worthwhile for future research to investigate whether these practices can be adopted more broadly and whether they should be included in the model. Further, a useful extension of our work would be to compare the practices of seasoned AP teachers to those who are newly teaching AP courses.

Our study validates some of the criticisms launched at the new AP curriculum. We have communicated our results to the Board's professional development team, who are working on improving their training of teachers for the new inquiry-based model. We also note that our findings generalize to teachers who are brand new to this curriculum and, as discussed above, have lower levels of training than science teachers nationally. More experienced teachers or those with master's degrees in STEM, even in less-resourced communities, likely face fewer obstacles to AP implementation. We hope that our research encourages more inquiry into the successes and challenges that teachers face in launching rigorous and student-directed AP instruction. The field would also benefit from more development of the logic model underlying AP courses, including greater specificity around which contextual and treatment components lead to specific student outcomes.

## Acknowledgments

This research was funded by the National Science Foundation (Award 1220092, Mark C. Long [PI], Dylan Conger [Co-PI], and Raymond McGhee Jr. [Co-PI]) and is registered in the American Economic Association's Registry for RCTs (ID 000140). We are grateful to our advisory committee members (Del Harnisch, Michal Kurlaender, Richard Murnane, Helen Quinn, and Aaron Rogat) for terrific guidance and feedback. Excellent research assistance was provided by Nicole Bateman, Kerry Beldoff, Grant H. Blume, Jordan Brown, Sarah Coffey, Bonnee Groover, Josette Arevalo Gross, Hernando Grueso Hurtado, Alec Kennedy, Jessica Mislevy, Kelsey Rote, Massiel Sepulveda, and Mariam Zameer. College Board staff provided answers to our questions about the Advanced Placement program and general feedback on the research design, but the College Board did not provide financial support and was otherwise not involved in the production of this research.

## Appendix A

**Table A1**

Context Indicators

Context	Context Indicator	Value Minimum	Value Maximum	Unadjusted Mean	Range Adjusted Mean (0–1)
Well-prepared and motivated teachers	Hold a teaching certificate	0 (No)	1 (Yes)	0.92	.92
	Undergraduate major in STEM	0 (No)	1 (Yes)	0.94	.94
	Single subject credential in science	0 (No)	1 (Yes)	0.67	.67
	Science coursework outside of major	0 (No)	1 (Yes)	0.64	.64
	Master's degree or higher	0 (No)	1 (Yes)	0.36	.36
	Previously taught AP/IB or honors course	0 (No)	1 (Yes)	0.83	.83
	Number of recent professional development trainings	0	5	3.10	.62

Context	Context Indicator	Value Minimum	Value Maximum	Unadjusted Mean	Range Adjusted Mean (0–1)
	Professional training helped course implementation	1 (Significant hindrance)	5 (Significant help)	4.05	.76
	Typical instruction involves student self-directed learning	1 (Not at all)	4 (A great deal)	3.13	.71
	Typical instruction involves hands on learning	1 (Not at all)	4 (A great deal)	3.42	.81
	Typical instruction involves class discussion and debates	1 (Not at all)	4 (A great deal)	2.60	.53
	Typical Instruction involves integration of technology and face-to-face discussion	1 (Not at all)	4 (A great deal)	3.00	.67
	Similarity between typical instruction and AP science instructional approach	1 (Not at all similar)	3 (Very similar)	2.03	.51
	Alignment of AP curriculum to typical instructional approach helped course implementation	1 (Significant hindrance)	5 (Significant help)	3.85	.71
	Attended summer institute training	0 (No)	1 (Yes)	0.77	.77
	Summer institute prepared me well	1 (Strongly disagree)	5 (Strongly agree)	3.40	.60
Well-prepared and motivated students	Student completed prerequisite science course <sup>a</sup>	0 (No)	1 (Yes)	0.84	.84
	Student completed prerequisite math course <sup>a</sup>	0 (No)	1 (Yes)	0.88	.88
	Student's grade point average <sup>a</sup> in prerequisite courses	0	4	3.14	.78
	Prior science knowledge / incoming ability of students helped course implementation	1 (Significant hindrance)	5 (Significant help)	2.98	.49
Sufficient planning time and resources	Planning time helped course implementation	1 (Significant hindrance)	5 (Significant help)	2.97	.49
	Resources and materials helped course implementation	1 (Significant hindrance)	5 (Significant help)	4.06	.76
Supportive district and school conditions	Principal or other school leaders helped course implementation	1 (Significant hindrance)	5 (Significant help)	3.29	.57
	Fellow teachers at school helped course implementation	1 (Significant hindrance)	5 (Significant help)	3.27	.57
	Teachers at other schools helped course implementation	1 (Significant hindrance)	5 (Significant help)	3.74	.69
	Frequent networking with teachers at other schools	1 (Never)	6 (Once a week or more)	3.18	.44
	Inquiry-based learning regularly discussed during team meetings	0 (No)	1 (Yes)	0.26	.26

Context	Context Indicator	Value Minimum	Value Maximum	Unadjusted Mean	Range Adjusted Mean (0–1)
	Science content or facts regularly discussed during team meetings	0 (No)	1 (Yes)	0.22	.22
	Have a mentor or coach at school that provides support	0 (No)	1 (Yes)	0.24	.24
	Alignment of AP curriculum and school’s curriculum helped course implementation	1 (Significant hindrance)	5 (Significant help)	3.26	.45
	Support from school/district leaders in curriculum development	1 (No support)	4 (Excellent support)	1.97	.24
	Support from school/district leaders in instructional materials	1 (No support)	4 (Excellent support)	2.26	.31
	Support from school/district leaders in pedagogy	1 (No support)	4 (Excellent support)	2.01	.25

*Note.* All indicators other than those marked with a superscripted *a* are measured from survey of Advanced Placement (AP) teachers. Middle values are as follows: 2 (Small hindrance), 3 (No impact), 4 (Small help); 2 (A little), 3 (Somewhat); 2 (Somewhat similar); 2 (Disagree), 3 (Neutral), 4 (Agree); 2 (Once during the year or less), 3 (A few times over the school year), 4 (Once a month), 5 (A few times a month); and 2 (Some, but not enough support), 3 (Adequate support). STEM = science, technology, engineering, and mathematics; IB = International Baccalaureate.

<sup>a</sup>Denotes indicators that were determined by the study team’s analysis of student transcripts.

**Table A2**

Treatment Indicators

Treatment	Treatment Indicator	Value Minimum	Value Maximum	Unadjusted Mean	Range Adjusted Mean (0–1)
College-level and inquiry-based syllabus	Syllabus audited by College Board (reviewed by college faculty)	0 (No)	1 (Yes)	0.96	.96
	Syllabus integrates disciplinary content and big ideas <sup>a</sup>	1 (Not present)	4 (Highly integrated)	2.81	.60
	Syllabus integrates practices and communications with content <sup>a</sup>	1 (Not present)	4 (Highly integrated)	2.38	.46
	Syllabus integrates practices <sup>a</sup> with hands-on activities	1 (Not present)	4 (Highly integrated)	2.75	.58
	Syllabus integrates issues around society, technology, and innovation with content <sup>a</sup>	1 (Not present)	3 (Integrated)	1.88	.44
Coverage of major science practices and learning objectives	Course organized around learning objectives	1 (Strongly disagree)	5 (Strongly agree)	4.22	.81
	Course organized around science practices	1 (Strongly disagree)	5 (Strongly agree)	3.87	.72
	Able to implement as laid out in syllabus	1 (Strongly disagree)	5 (Strongly agree)	3.24	.56
	Extent to which required content has been covered	1 (Very Little)	4 (Majority)	3.14	.71



Treatment	Treatment Indicator	Value Minimum	Value Maximum	Unadjusted Mean	Range Adjusted Mean (0–1)
Academically challenging curriculum	Academic challenge of AP course compared to honors course	1 (A lot less challenging)	5 (A lot more challenging)	4.67	.92
	Academic challenge of AP course compared to regular course	1 (A lot less challenging)	5 (A lot more challenging)	4.51	.88
Project-based and independent classroom activities	Students work independently	1 (Never)	5 (Very frequently)	3.55	.64
	Students work in a small group	1 (Never)	5 (Very frequently)	4.15	.79
	Students read book (non-textbook) or magazine about science	1 (Never)	5 (Very frequently)	2.42	.36
	Students are asked to apply their knowledge to solve a new problem without teacher guidance	1 (Never)	5 (Very frequently)	3.42	.61
	Students design their own projects or experiments	1 (Never)	5 (Very frequently)	3.05	.51
	Students engage in “hands-on learning” by using materials, tools, kits, and/or supplies	1 (Never)	5 (Very frequently)	3.67	.67
	Students present what they learned about a topic to the class	1 (Never)	5 (Very frequently)	2.94	.48
Integrated use of technology	Students used technology to practice concepts after they learned them in class	1 (Never)	5 (Very frequently)	3.43	.61
	Students used technology to conduct interactive simulations to explore relationships in data	1 (Never)	5 (Very frequently)	3.17	.54
	Students used technology to create graphical presentations of information	1 (Never)	5 (Very frequently)	3.13	.53
	Students used technology to develop collaborative projects and/or group presentations	1 (Never)	5 (Very frequently)	3.07	.52

*Note.* All indicators other than those marked with a superscripted *a* are measured from survey of Advanced Placement (AP) teachers. Middle values are as follows: 2 (Isolated), 3 (Partially integrated); 2 (Disagree), 3 (Neutral), 4 (Agree); 2 (Some), 3 (Most but not all); 2 (Somewhat less challenging), 3 (About as challenging), 4 (Somewhat more challenging); and 2 (Rarely), 3 (Occasionally), 4 (Frequently).

<sup>a</sup>Denotes indicators that were determined by the study team’s analysis of the AP teacher’s syllabus.

## Biography

MARK C. LONG is a professor of public policy and governance at the University of Washington, Evans School of Public Policy and Governance, Box 353055, Seattle, WA 98105; marklong@uw.edu. His research examines the effects of public policies on economic opportunity and efficient social mobility, with emphasis on estimating the benefits and costs of those policies.

DYLAN CONGER is a professor at The George Washington University, Trachtenberg School of Public Policy and Public Administration, 805 21st Street NW, Washington, DC 20052; dconger@gwu.edu. Her research focuses on explaining disparities in achievement between social groups and evaluating policies aimed at reducing those disparities.

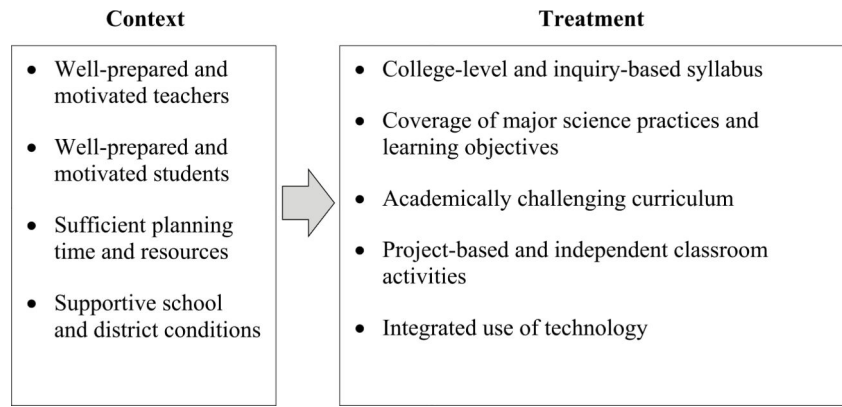
RAYMOND MCGHEE JR. is a senior director of evaluation at Equal Measure, a nonprofit services firm based in Philadelphia, 520 Walnut St., Suite 1450, Philadelphia, PA 19106; rmcghee@equalmeasure.org. His research and evaluation work examines programs and policies designed to broaden participation of underserved populations in STEM education at the secondary and postsecondary levels, as well as the workforce.

## REFERENCES

- Abdulkadiro lu A, Pathak PA, & Walters CR (2018). Free to choose: Can school choice reduce student achievement? *American Economic Journal: Applied Economics*, 10(1), 175–206.
- Adelman C (2006). *The Toolbox revisited: Paths to degree completion from high school through college*. Washington, DC: U.S. Department of Education.
- Angrist JD, Imbens GW, & Rubin DB (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, 91(434), 444–455.
- Angrist JD, & Pischke J (2009). *Mostly harmless econometrics*. Princeton, NJ: Princeton University Press.
- Avery C, Gurantz O, Hurwitz M, & Smith J (2017). Shifting college majors in response to Advanced Placement exam scores. *Journal of Human Resources*, 53(4), 918–956.
- Bloom HS, Orr LL, Bell SH, Cave G, Doolittle F, Lin W, & Bos JM (1997). The benefits and costs of JTPA Title II-A programs: Key findings from the National Job Training Partnership Act Study. *Journal of Human Resources*, 32(3), 549–576.
- Boatman A, & Long BT (2018). Does remediation work for all students? How the effects of postsecondary remedial and developmental courses vary by level of academic preparation. *Educational Evaluation and Policy Analysis*, 40(1), 29–58.
- Bowie L (2013, 8 17). Maryland schools have been leader in Advanced Placement, but results are mixed. *The Baltimore Sun* Retrieved from <http://www.baltimoresun.com/news/maryland/bsmd-advanced-placement-classes-20130817-story.html>
- College Board. (2011a). *AP Biology: Curriculum framework 2012–2013*. New York: Author.
- College Board. (2011b). *AP Chemistry: Curriculum framework 2013–2014*. New York: Author.
- College Board (2014). *AP Chemistry: Course and exam description revised edition effective Fall 2014*. New York: Author.
- College Board (2015). *AP Biology: Course and exam description revised edition effective Fall 2015*. New York: Author.
- College Board. (2017). *AP course and exam redesign*. Retrieved from <https://aphighered.collegeboard.org/courses-exams/course-exam-resdeign>
- College Board. (2019a). *All in*. Retrieved from <https://professionals.collegeboard.org/k-12/all-in>
- College Board. (2019b). *AP for all*. Retrieved from <https://professionals.collegeboard.org/testing/states-local-governments/partnerships/nyc/ap-all>
- Conger D, Kennedy AL, Long MC, & McGhee R Jr. (in press). The effect of Advanced Placement science on students' skills, confidence and stress. *Journal of Human Resources*.
- Conger D, Long MC, & Iatarola P (2009). Explaining race, poverty, and gender disparities in advanced coursetaking. *Journal of Policy Analysis and Management*, 28(4), 555–576.
- Education Commission of the States (ECS). (2016). *50-state comparison: Advanced Placement policies*. Retrieved from <https://www.ecs.org/advanced-placement-policies>
- Farkas S, & Duffett A (2009). *Growing pains in the Advanced Placement program: Do tough trade offs lie ahead?* Washington, DC: Thomas B. Fordham Institute.

- General education in school and college: A committee report by members of the faculties of Andover, Exeter, Lawrenceville, Harvard, Princeton, and Yale. (1952). Cambridge, MA: Harvard University Press.
- Holstead MS, Spradlin TE, McGillivray ME, & Burroughs N (2010). The impact of Advanced Placement incentive programs (Education Policy Brief, 8[1]). Bloomington, IN: Center for Evaluation and Education Policy.
- Hull DM, Hinerman KM, Ferguson SL, Chen Q, & Näslund-Hadley EI (2018). Teacher-led math inquiry: A cluster randomized trial in Belize. *Educational Evaluation and Policy Analysis*, 40(3), 336–358.
- Iatarola P, Conger D, & Long MC (2011). Determinants of high schools' advanced course offerings. *Educational Evaluation & Policy Analysis*, 33(3), 340–359.
- International Baccalaureate Organization. (2016). The IB Diploma Programme statistical bulletin. Grand-Saconnex, Switzerland: Author.
- Imbens GW, & Angrist JD (1994). Identification and estimation of local average treatment effects. *Econometrica*, 62(2), 467–475.
- Jackson K (2010). A little now for a lot later: An evaluation of the Texas Advanced Placement incentive program. *Journal of Human Resources*, 45(3), 591–639.
- Jackson K (2014). Do college-prep programs improve long-run outcomes? *Economic Inquiry*, 52(1), 72–99.
- Judson E (2017). Science and mathematics Advanced Placement exams: Growth and achievement over time. *Journal of Educational Research*, 110, 209–217.
- Judson E, & Hobson A (2015). Growth and achievement trends of Advanced Placement (AP) exams in American high schools. *American Secondary Education*, 43(2), 59–76.
- Klopfenstein K (2004). Advanced Placement: Do minorities have equal opportunity? *Economics of Education Review*, 23(2), 115–131.
- Klopfenstein K, & Thomas MK (2009). The link between Advanced Placement experience and early college success. *Southern Economic Journal*, 75(3), 873–891.
- Kurth LA, Anderson CW, & Palincsar AS (2002). The case of Carla: Dilemmas of helping all students to understand science. *Science Education*, 86(3), 287–313.
- Lichten W (2010). Whither Advanced Placement-now? In Sadler PM, Sonnert G, Tai RH, & Klopfenstein K (Eds.), *AP: A critical examination of the Advanced Placement program* (pp. 233–234). Cambridge, MA: Harvard Education Press.
- Litzler E, Samuelson CC, & Lorah JA (2014). Breaking it down: Engineering student STEM confidence at the intersection of race/ ethnicity and gender. *Research in Higher Education*, 55(8), 810–832.
- Malkus N (2016). *AP at scale*. Washington, DC: American Enterprise Institute.
- Monk DH, & Haller EJ (1993). Predictors of high school academic course offerings: The role of school size. *American Educational Research Journal*, 30, 3–21.
- National Research Council. (2002). *Learning and understanding: Improving advanced study of mathematics and science in US high schools*. Washington, DC: The National Academies Press.
- National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas* Committee on a Conceptual Framework for New K-12 Science Standards. Board on Science Education, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.
- Reardon SF, Ho AD, Shear BR, Fahle EM, Kalogrides D, & DiSalvo R (2018). Stanford education data archive (Version 2.1). Retrieved from <http://purl.stanford.edu/db586ns4974>
- Rodriguez A, & McGuire KM (2019). More classes, more access? Understanding the effects of course offerings on Black-White gaps in Advanced Placement course-taking. *Review of Higher Education*, 42(2), 641–679.
- Schneider J (2013, 3 9). What the AP program can't do. *Washington Post Blog*.
- Smith J, Hurwitz M, & Avery C (2017). Giving college credit where it is due: Advanced Placement exam scores and college outcomes. *Journal of Labor Economics*, 35(1), 67–147.

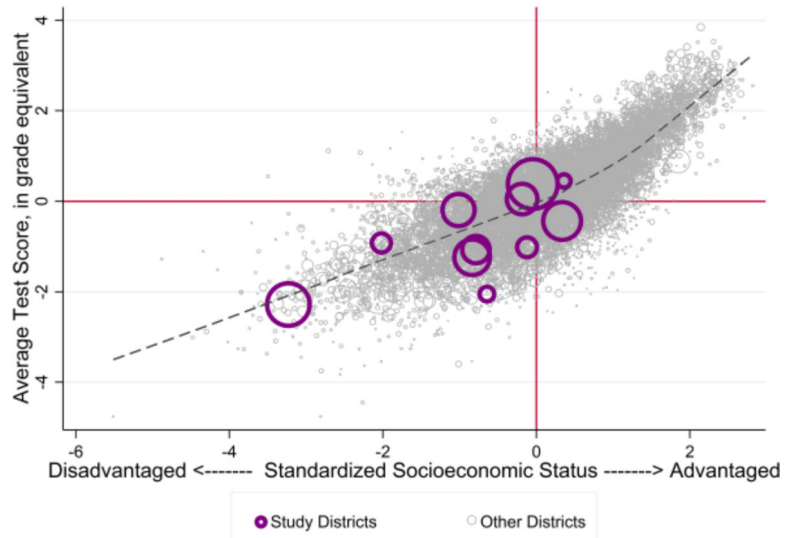
- Staiger D, & Stock JH (1997). Instrumental variables regression with weak instruments. *Econometrica*, 65(3), 557–586.
- Tai RH (2008). Posing tougher questions about the Advanced Placement program. *Liberal Education*, 94(3), 38–43.
- Thomas N, Marken S, Gray L, & Lewis L (2013). Dual credit and exam-based courses in US public high schools: 2010–11. First look (NCES 2013–001). Washington, DC: National Center for Education Statistics.
- Tierney J (2012, 10 13). AP classes are a scam. *The Atlantic*.
- U.S. Department of Education, National Center for Education Statistics. (2017). Public high school 4-year adjusted cohort graduation rate (ACGR), by race/ethnicity and selected demographic characteristics for the United States, the 50 states, and the District of Columbia: School year 2015–16. Retrieved from [https://nces.ed.gov/ccd/tables/ACGR\\_RE\\_and\\_characteristics\\_2015-16.asp](https://nces.ed.gov/ccd/tables/ACGR_RE_and_characteristics_2015-16.asp)
- U.S. Department of Education, National Center for Education Statistics, Schools and Staffing Survey, Public School Teacher and BIE School Teacher Data Files, 2007–08. (n.d.). Number and percentage distribution of public school teachers, by reported level of agreement with the statement, ‘The school administration’s behavior toward the staff is supportive and encouraging’ and selected school characteristics: 2007–08. Retrieved from [http://nces.ed.gov/surveys/sass/tables/sass0708\\_013\\_t1n.asp](http://nces.ed.gov/surveys/sass/tables/sass0708_013_t1n.asp)
- Warne RT, Sonnert G, & Sadler PM (2019). The relationship between Advanced Placement mathematics courses and students’ STEM career interest. *Educational Researcher*, 48(2), 101–111.
- The White House, Domestic Policy Council and Office of Science and Technology Policy. (2006). American Competitiveness Initiative. Retrieved from <https://georgewbush-whitehouse.archives.gov/stateoftheunion/2006/aci/aci06-booklet.pdf>
- Xu D, Solanki S, McPartlan P, & Sato B (2018). EASEing students into college: The impact of multidimensional support for underprepared students. *Educational Researcher*, 47(7), 435–450.



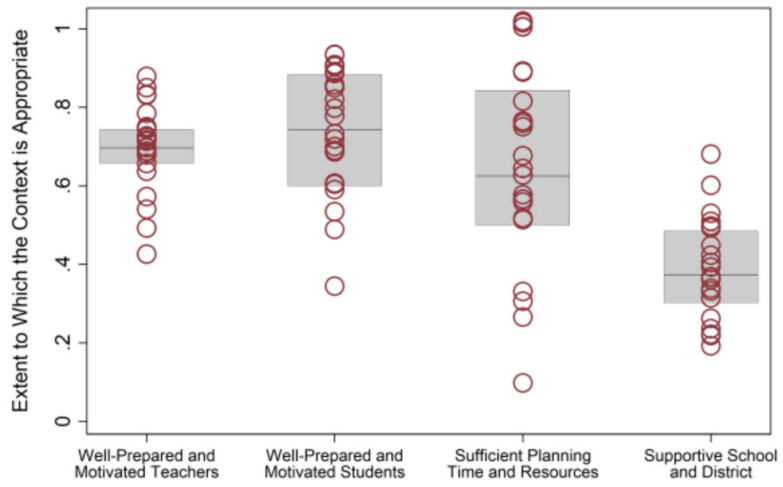
**FIGURE 1.**  
Key context and treatment components of a well-implemented AP science course.



**FIGURE 2.**  
Map with all the sites.



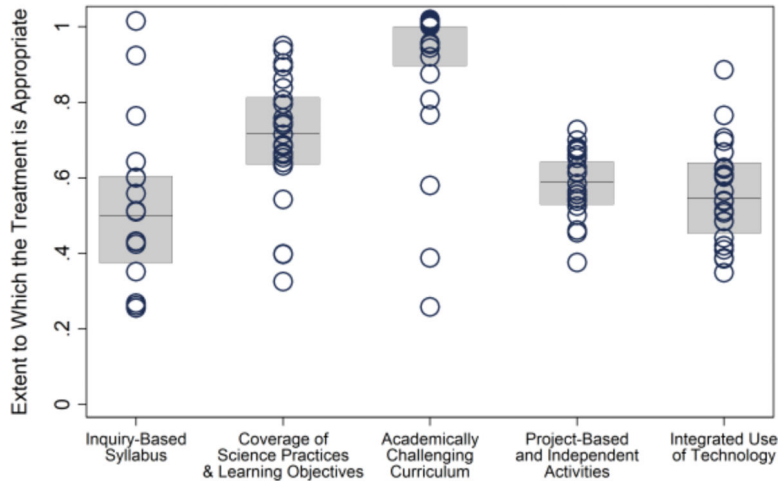
**FIGURE 3. Participating districts' neighborhood socioeconomic status and school test scores.**  
*Note.* Data from Reardon, Kalogrides, and Shores (2016). Each circle represents one school district in the United States. X-axis is the standardized socioeconomic status of the district's neighborhood defined as the first principal component factor score based on measures of median income, percent with a bachelor's degree or higher, poverty rate, Supplemental Nutrition Assistance Program (SNAP) rate, single-motherheaded household rate, and unemployment rate. Y-axis is the district's average test score, in grade equivalents, based on the averaged spring math and English scores for students in Grades 3–8 for 2009–2013, with the expected level of achievement standardized to zero. The size of each circle is proportional to the district's enrollment. The dashed line is a lowess curve created using Stata's default settings and roughly shows the predicted test score as a function of the neighborhood's SES.



**FIGURE 4. Evaluation of context.**

*Note.* Each marker represents one school. The height of the marker is slightly jittered to create separation. Interquartile range is illustrated by gray bars.





**FIGURE 5. Evaluation of treatment.**

*Note.* Each marker represents one school. The height of the marker is slightly jittered to create separation. Interquartile range is illustrated by gray bars.

**Table 1**

Participating School and Teacher Descriptives

<b>Panel A: Schools</b>	<b>Participating</b>	<b>Others</b>
Average enrollment	1,409	723
Free or reduced-price lunch	.700	.438
Asian	.055	.050
Black	.349	.154
Hispanic	.410	.221
White	.164	.537
Adjusted cohort graduation rate	.843	.802
District's instruction expenditures per pupil	\$6,561	\$5,636
District's student services expenditures per pupil	\$3,787	\$3,385
<b>Panel B: Schools</b>	<b>Participating</b>	<b>Others</b>
Age: Under 30	.407	.160
Age: 30–49	.432	.553
Age: 50 or over	.161	.287
Female	.630	.536
Hispanic or Latino	.111	.051
Race: American Indian or Alaska Native	.000	.009
Race: Asian American	.111	.041
Race: Black	.111	.060
Race: Native Hawaiian or other Pacific Islander	.000	.004
Race: White	.778	.896
Years of experience 2	.290	.085
Years of experience 5	.481	.234
Hold a teaching certificate	.926	.945
Undergraduate major in science, technology, engineering, and mathematics (STEM)	.944	.747
Single subject credential in science	.630	.823
Master's degree or higher	.356	.615

*Note.* Panel A source is the 2013–14 Common Core Data. “Others” in Panel A refers to other public high schools in the U.S. Adjusted cohort graduation rate is the percentage of the students in a ninth-grade cohort who graduate within 4 years (U.S. Department of Education, 2017). Panel B source is the Teacher Survey (N = 27) and 2011–12 Schools and Staffing Survey. “Others” in Panel B refers to public and private high school teachers in the U.S. High school science teachers are defined as teachers of Grades 9–12 whose main teaching assignment is in the natural sciences.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 2**

Student Reports on Advanced Placement (AP) Science Courses

	Control Group	Complier Mean	Local Average Treatment Effect ( $\beta$ )	Sig.
Academically challenging curriculum		.68	.37 (.07)	***
Course was intellectually challenging <sup>a</sup>				
Often receive homework <sup>b</sup>		.47	.39 (.10)	***
Teacher set high standards <sup>a</sup>		.74	.19 (.07)	**
Students were driven to succeed <sup>a</sup>		.56	.24 (.10)	**
Participate in hands-on learning <sup>b</sup>		.52	.25 (.11)	**
Design own projects or experiments <sup>b</sup>		.24	.18 (.09)	*
Work in small groups <sup>b</sup>		.54	.17 (.07)	**
Work independently <sup>b</sup>		.59	-.02 (.08)	
Present what you learned <sup>b</sup>		.32	-.05 (.10)	
Apply knowledge to solve new problem <sup>b</sup>		.48	.08 (.08)	
Use of integrated technology		.26	.05 (.08)	
Conduct interactive simulations <sup>b</sup>		.27	.05 (.07)	
Create graphical presentations <sup>b</sup>		.34	.06 (.08)	
Develop collaborative projects <sup>b</sup>		.28	.07 (.07)	
Practice concepts <sup>b</sup>				

Note. Number of Observations = 1,417. All outcomes are binary. Results are weighted by the inverse probability of completing the survey. Standard errors clustered by School  $\times$  Cohort are in parentheses.

<sup>a</sup> 1 if strongly agree or agree versus 0 if neutral, disagree, or strongly disagree.

<sup>b</sup> 1 if very frequently or frequently versus 0 if occasionally, rarely, or never.

\* Two-tailed statistics significance at the 10% level.

\*\* Two-tailed statistics significance at the 5% level.

\*\*\* Two-tailed statistical significance at the 1% level.