

Advanced Placement® Exam-Taking and Performance:

Relationships with First-Year Subject Area
College Grades

By Brian F. Patterson, Sheryl Packman, and Jennifer L. Koblin



Brian F. Patterson is an assistant research scientist at the College Board.

Sheryl Packman was an associate research scientist at the College Board.

Jennifer L. Kobrin is a research scientist at the College Board.

Acknowledgments

The authors would like to acknowledge the helpful feedback of Vytas Laitusis and Ann A. O’Connell on earlier drafts of this manuscript and Michael Chajewski for his review of the tables. James Monk and Keith Cothrun also provided thoughtful advice on analyses of the world language student data.

About the College Board

The College Board is a mission-driven not-for-profit organization that connects students to college success and opportunity. Founded in 1900, the College Board was created to expand access to higher education. Today, the membership association is made up of more than 5,900 of the world’s leading educational institutions and is dedicated to promoting excellence and equity in education. Each year, the College Board helps more than seven million students prepare for a successful transition to college through programs and services in college readiness and college success — including the SAT® and the Advanced Placement Program®. The organization also serves the education community through research and advocacy on behalf of students, educators and schools.

For further information, visit www.collegeboard.org.

© 2011 The College Board. College Board, Advanced Placement Program, AP, SAT and the acorn logo are registered trademarks of the College Board. inspiring minds is a trademark owned by the College Board.

**Visit the College Board on the Web:
www.collegeboard.org.**

VALIDITY

Contents

Executive Summary	4
Introduction	4
Method	7
Sample	7
Measures	7
Analyses	9
Results	10
Descriptive Statistics	10
Cross-Classified Multilevel Modeling	13
Mathematics Course Performance	14
Computer Science Course Performance	16
Engineering Course Performance	18
Natural Science Course Performance	20
Social Science Course Performance	22
History Course Performance	24
English Course Performance	26
World Language Course Performance	28
Art and Music Course Performance	30
Discussion	32
Limitations and Directions for Future Research	34
Conclusions	34
References	35

Tables

1.	AP [®] Examinees by Subject Area Sample.....	11
2.	Student Characteristics by Subject Area Sample.....	37
3.	High School Characteristics Weighted by Students in Each Subject Area Sample.....	38
4.	College Characteristics Weighted by Students in Each Subject Area Sample.....	39
5.	Cross-Classified Multilevel Model Covariance Parameter Estimates for Null Models.....	13
6.	Means (Standard Deviations) on Academic Characteristics of Mathematics Sample.....	14
7.	Mathematics Cross-Classified Multilevel Model Results.....	15
8.	Means (Standard Deviations) on Academic Characteristics of Computer Science Sample.....	16
9.	Computer Science Cross-Classified Multilevel Model Results.....	17
10.	Means (Standard Deviations) on Academic Characteristics of Engineering Sample.....	18
11.	Engineering Cross-Classified Multilevel Model Results.....	19
12.	Means (Standard Deviations) on Academic Characteristics of Natural Science Sample.....	20
13.	Natural Science Cross-Classified Multilevel Model Results.....	21
14.	Means (Standard Deviations) on Academic Characteristics of Social Science Sample.....	22
15.	Social Science Cross-Classified Multilevel Model Results.....	23
16.	Means (Standard Deviations) on Academic Characteristics of History Sample.....	24
17.	History Cross-Classified Multilevel Model Results.....	25
18.	Means (Standard Deviations) on Academic Characteristics of English Sample.....	26

19.	English Cross-Classified Multilevel Model Results	27
20.	Means (Standard Deviations) on Academic Characteristics of World Language Sample	28
21.	World Language Cross-Classified Multilevel Model Results	29
22.	Means (Standard Deviations) on Academic Characteristics of Art and Music Sample	30
23.	Art and Music Cross-Classified Multilevel Model Results	31

Executive Summary

The purpose of this study was to examine the effects of Advanced Placement® (AP®) exam participation and performance on college grades for courses taken in the same subject area as students' AP Exam(s). Students' first-year college subject area grade point averages (SGPAs) were examined in nine subject areas: mathematics, computer science, engineering, natural science, social science, history, English, world language, and art and music. Using cross-classified multilevel modeling for each subject area separately, and controlling for gender, racial or ethnic identity, socioeconomic status and prior academic ability, as average AP Exam score in each subject area increased, expected SGPA increased. Model selection procedures led us to include the number of AP Exams taken in the model for only three of the nine subject areas: engineering, natural science, and social science.

Introduction

Exploration of the plethora of possible career opportunities often gains momentum in high school when students are first able to exert their preferences in course selection. This most often occurs after students have attained their fundamental education requirements and are able to enroll in advanced academic courses that spark their interests. Through enrollment in Advanced Placement® (AP®), International Baccalaureate, dual enrollment¹ and other college-level courses, students can also use this opportunity as a way to jump-start their college education.

AP is a program of college-level courses and examinations offered by the College Board. The content and skills taught in each AP course intend to represent the content and skills taught in a corresponding introductory college-level course. As such, students who earn a particular grade on an AP Exam, depending on the individual credit policy of the institution, can place out of, or obtain credit for, the corresponding introductory college course. Since the inception of AP in 1955, the program has gone through many changes in terms of the number of courses and exams offered. As of the 2009-10 academic year, there were 33 courses and exams in a variety of subject areas.

Given that the primary motivation for taking an AP Exam is the opportunity to earn credit for an introductory college course or courses or to place into advanced course work, past research on AP focused mainly on the validity of AP Exam scores for predicting student performance in the subsequent course, that is, the course(s) students would take immediately after completing the introductory course. Studies in this area most often show that students who perform well on the AP Exam and place out of the introductory college course perform as well or better in the subsequent course as students who take the introductory course (Casserly, 1986; Dodd, Fitzpatrick, De Ayala, & Jennings, 2002; Hargrove, Godin, & Dodd, 2008; Keng & Dodd, 2008; Morgan & Crone, 1993; Morgan & Ramist, 1998; Willingham & Morris, 1986).

Morgan and Crone (1993) aver that the success of the AP program should be measured not only by performance differences between AP and non-AP students in subsequent course work, but also by differences in the future interest AP students show in their AP Exam subjects and college performance in courses related to their AP Exam subjects. They compared college freshmen in 1988 and 1989 who took an AP Exam in Biology, Calculus AB, or Chemistry to non-AP students to examine whether there were differences in the amount and types of college courses taken as well as the grade in the subsequent course (due to placement out of the introductory course from their AP grade). Results showed that the AP students who took Calculus AB or Chemistry continued to pursue a course of study in these areas. In addition, students who received an AP grade of 3 or higher received higher mean grades in the

subsequent course than did non-AP students. Finally, students who received an AP grade of 4 or 5 did extremely well in their initial college course work in their AP subject. Additional research expands these findings to show that students who take AP Exams earn more college credits than non-AP students in their AP subjects (Morgan & Maneckshana, 2000; Morgan & Ramist, 1998).

Willingham and Morris (1986) tracked the college academic pathways of over 1,000 AP students who attended private liberal arts colleges for a period of four years. To study the effects of AP Exam performance on subsequent performance in college they grouped the exams into six subject areas: English, history, foreign language, biological sciences, mathematics, and physical sciences. Similar to other studies, the researchers found that AP students were more likely to enroll in courses in the same AP subject area and earn higher grades in those courses than non-AP students. Overall, the AP students outperformed the group of non-AP students who were matched on academic ability. In addition, students with AP grades of 4 or 5 earned much higher college GPAs than would have been predicted by typical measures of college readiness (i.e., SAT® scores and high school rank). This study did not find any additional differences in terms of college performance between the AP and non-AP groups based on gender, high school affiliation (i.e., public or private), or college selectivity.

More recent studies have also looked at performance in college courses in the AP subject area. For example, Dodd, Fitzpatrick, De Ayala, and Jennings (2002) examined AP performance for students entering a large public university in Texas from 1995–1999. Results showed that AP students who received AP credit (i.e., a grade of 3 or higher) for the English Language and Composition, Calculus AB, or Biology exams attained the same or higher grades in subsequent courses and achieved higher GPAs in the subject area of the AP Exam than AP students who did not earn credit and took the introductory course, non-AP students of similar ability, and dual enrollment students.

Similarly, Morgan and Klaric (2007) compared the college academic records of students who did and did not take AP Exams. Their sample included over 72,000 students from the class of 1994 who attended 27 institutions. The institutions in their sample were targeted because they received large volumes of AP score reports in connection with the undergraduate admissions process. Students were tracked for five years on course-taking behavior and performance in subject areas in which an AP Exam was taken. Results showed that AP students performed well in the course that followed the introductory course for which they received advanced placement or credit. With the exception of AP English Literature, compared to their non-AP counterparts, AP students took a greater number of college courses in their AP Exam subjects. They also found that AP students outperformed their non-AP counterparts in the subsequent course and that there was a positive relationship between AP Exam grades and college course grades.

In another study, Keng and Dodd (2008) analyzed AP performance on 10 exams with subsequent college performance. Their sample included first-year students who attended a large public institution in Texas from 1998–2001. They compared performance for AP and non-AP students who were matched on ability. Results showed that AP students consistently outperformed non-AP students in total first-year GPA and earned higher GPAs in each of the AP subject areas.

Hargrove, Godin, and Dodd (2008) also utilized data from Texas; however, their sample included students who graduated from high school in Texas from 1998–2002, took at least one of seven AP Exams (i.e., English Language and Composition, English Literature, Calculus AB, Biology, Chemistry, U.S. History, and Spanish Language), and attended any public institution in Texas.

Additional predictor variables included gender and racial or ethnic identity as well as SAT scores and free or reduced price lunch status, which were included to control for college readiness and socioeconomic status, respectively. Outcome variables included first- and fourth-year GPA, credit hours earned, and graduation within four years. When the sample was aggregated across exams, AP students who took both the course and exam outperformed students who: only took the AP course; only took the AP Exam; took neither the course nor an exam (non-AP students); and non-AP dual enrollment students on GPA, graduation rate, and overall number of credit hours. While there were differences in college performance by gender and racial or ethnic identity for the specific AP Exam taken, the overall finding that AP students who took the course and exam outperformed their counterparts was consistent across subgroups.

The purpose of this study was to examine the effects of AP Exam participation and performance on college grades in nine subject areas, as compared with non-AP examinees.

The purpose of this study was to examine the effects of AP Exam participation and performance on college grades in nine subject areas, as compared with non-AP examinees. Overall, nine subject areas were examined: mathematics, computer science, engineering, natural science, social science, history, English, world language, and art and music. In accordance with past research (e.g., Dodd, Fitzpatrick, De Ayala & Jennings, 2002; Keng & Dodd, 2008), it was hypothesized that with either greater numbers of subject area AP Exams taken or higher mean grades on those subject area AP Exams, students would earn higher subject area first-year grade point averages (SGPAs).

Through the design of this study, some key methodological limitations of past studies were mitigated. First, by including outcome data on students who attended more than 100 four-year colleges and universities that varied on a variety of characteristics, this study had greater generalizability to the entire population of college students than studies that focused on data from one institution (Dodd, Fitzpatrick, De Ayala, & Jennings, 2002; Keng & Dodd, 2008; Morgan & Crone, 1993) or a group of homogenous institutions (Willingham & Morris, 1986). Second, by utilizing cross-classified multilevel modeling to examine student effects and controlling for the random variation

of SGPAs across the high schools and colleges attended, the results of this study addressed potential confounds of past research (Dodd, Fitzpatrick, De Ayala & Jennings, 2002; Keng & Dodd, 2008; Morgan & Crone, 1993; Morgan & Klaric, 2007). Given that the outcome of interest (SGPA) was observed within colleges and because variation in that outcome was likely to be more easily attributed to differences across colleges — rather than across high schools — more random variation in student performance was expected to be associated with the college, than with the high school. Similarly, this study controlled for prior academic achievement (i.e., SAT scores and high school grade point average) as well as student (i.e., racial or ethnic identity, gender, and parental education), high school (i.e., location, mean enrollment per grade, and affiliation), and college (control, admittance rate, and undergraduate enrollment) demographic characteristics that have been related to academic success (Kobrin, Patterson, Shaw, Mattern, & Barbuti, 2008; Mattern, Patterson, Shaw, Kobrin, & Barbuti, 2008) and excluded from other

studies (e.g., Keng & Dodd, 2008). Finally, this study analyzed data from lower volume AP Exams (e.g., AP Computer Science A) that were typically excluded from prior research due to a lack of data (e.g., Dodd, Fitzpatrick, De Ayala, & Jennings, 2002; Morgan & Klaric, 2007; Keng & Dodd, 2008; and Hargrove, Godin, & Dodd, 2008).

Method

Sample

The sample included 110 four-year colleges and universities that were originally recruited as part of a national study to evaluate the validity of the SAT after its revision in March 2005. Recruitment plans were based on a stratified sampling scheme of four-year colleges and universities that took into account U.S. region, size, selectivity (as measured by admittance rate) and control (i.e., public or private). Participating institutions provided all first-year course grades for their entire cohort of first-time, first-year students entering in the fall of 2006. A nominal stipend was offered to offset the resources needed to create the student data file. See Kobrin, et al. (2008) for more information on the target population and sample characteristics.

After combining the college-level data with College Board data (i.e., SAT and AP Exam data), the sample included 195,099 first-time, first-year students who enrolled in college in the fall of 2006. Inclusion in one of the nine subject area samples, a student must have received a grade and credit for at least one college course in the subject area. When limited to those students with complete data for all three SAT sections, high school grade point average (HSGPA), high school characteristics, and credit and a grade in at least one course in at least one of the nine subject areas, the sample was reduced to 147,051. Students who transferred out of one of the 110 colleges but who had complete data on the required elements listed above were included in the analyses. Students who transferred into one of the sampled colleges during the 2006-07 academic year, however, were not included. High school affiliation was determined by the most recent high school attended. Table 1 presents the number of college students and the number of AP examinees who were included in each of the nine subject area samples.

Measures

AP Exam Participation and Performance. The data set used in this study could not track students who took an AP course, but not the corresponding AP Exam. Therefore, AP participation was operationalized as the number of AP Exams taken in the subject area. During the time students in this sample attended high school (2002-06), there were at least² 34 AP Exams offered in 20 subjects. AP performance was operationalized as the mean of AP Exam grades earned in the relevant subject area. The AP grading scale associates each grade from 1 to 5 with a student's ability to perform well in the corresponding subsequent college course

After combining the college-level data with College Board data (i.e., SAT and AP Exam data), the sample included 195,099 first-time, first-year students who enrolled in college in the fall of 2006.

(i.e., 1 = *no recommendation*; 2 = *possibly qualified*; 3 = *qualified*; 4 = *well qualified*; and 5 = *extremely well qualified*). We viewed the score scale as a set of categorical judgments of students' knowledge and skills and chose not to assume a linear relationship of the scale score with the SGPA. As such, mean AP Exam performance was treated as a categorical measure and students' mean AP Exam grades in each subject area were computed and then rounded to the nearest score on the 1 to 5 scale.

Subject Area First-Year College Performance. The participating colleges submitted course-taking data on each of their students that included the course label (e.g., ENGL 101); name (e.g., Freshman Composition); grade (e.g., 3.67); credit (e.g., 3.0); and the term (i.e., semester, trimester, quarter; e.g., second semester) in which the course was taken. Course grades generally ranged from 0.00 to 4.00, but five colleges granted course grades up to 4.30. Each course was coded into one of 14 subject areas (e.g., social science). The nine subject areas most compatible with existing AP Exam subject areas included mathematics, computer science, engineering, natural science, social science, history, English, world language, and art and music. The subject area first-year GPA (SGPA) was computed as the mean course grade within a subject area, weighted by the number of credits for each course. There was no limit placed on how many courses or credits a student would need to be included in the analyses.

The world language SGPA analyses were based on all courses where a non-English language (e.g., Spanish, Japanese) was taught. While alternate approaches would be to conduct a separate analysis either by each language family (e.g., the romance languages) or by each language (e.g., French), this study combined the data from all world languages. We propose that the grouping of all world languages is appropriate because of a phenomenon known as *language layering*, whereby students accrue general skills while learning a second language that are transferrable and assist in learning a third language (New Jersey Department of Education, 1999).

Student Characteristics. Self-reported gender, racial or ethnic identity, and parental education levels from the SAT Questionnaire were included in the cross-classified multilevel models to control for demographic and socioeconomic differences across students.

Overall Academic Preparedness. Two measures of prior academic preparedness were used in the current study: SAT scores and high school grade point average (HSGPA). The SAT consists of three sections: critical reading (SAT-CR), mathematics (SAT-M), and writing (SAT-W), with each section scaled from 200 to 800. Students' latest SAT scores were used in the analyses in order to measure academic preparedness as close as possible to the date of entrance into college. HSGPA was a student self-reported measure taken from the SAT Questionnaire, which was administered during SAT registration. Past research on this sample showed a high correlation between self-reported HSGPA and true HSGPA (Shaw & Mattern, 2010). HSGPA was reported on a 12-point scale, ranging from 0.00 to 4.33.

High School Characteristics. A variety of characteristics were included to characterize the high schools attended by students in each subject area sample. Most high school level variables were obtained from data prepared in the spring of 2006 by Market Data Retrieval, Inc. (MDR, 2006). Variables from this data source included the urbanicity of the area surrounding the high school (i.e., rural, suburban, or urban), the mean student enrollment per grade, and the high school affiliation (i.e., public, private Catholic, or private non-Catholic).

In addition to the MDR data (2006), the number of AP courses offered by each high school was included. Since data on the number of AP courses offered at each high school were not readily available for the time frame of the cohort's high school enrollment, this number was approximated. If at least five students took a particular AP Exam in a particular year, then the high

school was assumed to have offered that AP course in that year. The number of identified AP courses was then aggregated across the four years that this cohort attended high school (2003-06).

College/University Characteristics. Descriptive variables about the colleges and universities attended included control (i.e., public or private), admittance rate, and total undergraduate enrollment. These data were based on the Annual Survey of Colleges — which the College Board conducts to inform its annual college handbooks (College Board, 2007) — and were included to demonstrate the variation in size, selectivity, and control of the institutions in the sample.

Analyses

Once subject area samples were created, descriptive statistics were run to determine whether there were differences in composition between the samples based on student, high school, and college characteristic variables. In addition, the SGPA, HSGPA, and SAT scores were summarized for each subject area overall, by HSGPA category and by mean AP Exam grade category, to illustrate the performance differences on many of the measures that may be directly observed. HSGPA categories were created as follows: HSGPAs of 0.00, 1.00, and 1.33 mapped to the “D or F” category; 1.67, 2.00, and 2.33 to “C- to C+”; 2.67, 3.00, and 3.33 mapped to “B- to B+”; and 3.67, 4.00, and 4.33 mapped to “A- to A+.”

Given that some of the variability in students’ SGPAs may be attributable to the high school attended and/or the college attended, a cross-classified multilevel modeling approach was employed. The data were cross-classified, rather than nested hierarchically (e.g., high schools within school districts); because students from the same high school generally attended a variety of different colleges and each college admitted students from a variety of different high schools. By using cross-classification, the idiosyncrasies of the particular set of high schools and colleges in the sample were separated from the variables of interest (i.e., those relating to students’ participation in and performance on AP Exams) through the inclusion of random effect terms. The use of random intercept effects strengthened the ability to make inferences to the general population, because the fluctuations in SGPA due to the high school or college attended that cannot be explained by the student-level variables included in the model were separated out from the parameter estimates of the model to be estimated.

In this study, the intercept term represents the (expected) mean SGPA for students from the reference group (described below), attending each observed combination of high school and college. The expected mean for all levels of the demographic and academic variables is then modeled with random main effects for the high school and the college attended and their random interaction assumed zero, following Raudenbush & Bryk (2002; p. 378). Specifically, the models allowed students who were comparable on overall prior academic preparedness to have different predicted SGPA values, if students from a given high school consistently over- or under-performed

Given that some of the variability in students’ SGPAs may be attributable to the high school attended and/or the college attended, a cross-classified multilevel modeling approach was employed.

relative to the expected SGPA under that model. For example, consider two students with equally high latent mathematics ability; one attended a high school with strict grading standards, while the other attended a high school with less rigorous grading standards. In this case, failing to include the random effect for high school may have resulted in estimates of mathematics SGPA that were biased. For more information on cross-classified multilevel modeling, see Beretvas (2008).

In the cross-classified multilevel model analyses, four variables (SAT-CR, SAT-M, SAT-W, and HSGPA) were grand-mean centered to the sample of 147,051 students who took the SAT, reported their HSGPA, and for whom high school data was available. This approach was taken so that the reference group for each model would be more meaningful: White males whose parents' highest education level was a bachelor's degree (and whose best spoken language was English alone, for the world language analysis), who took none of the subject area AP Exams, and whose SAT-CR, SAT-M, SAT-W, and HSGPA scores were equal to the grand mean. In addition, for the purposes of the cross-classified multilevel model, SAT scores were divided by 100 so that the scale more closely matched that of HSGPA and SGPA.

Four models were estimated to examine variability and effects of AP Exam participation and performance for each of the nine subject areas. All predictor variables were included at the student level. Model 1 included gender, racial or ethnic identity, highest parental education level, SAT-CR, SAT-M, SAT-W, and HSGPA and served as the comparison for the three remaining models. Models 2 and 3 added the number of subject area AP Exams and average subject area AP Exam score, respectively, to Model 1 and Model 4 contained both of these AP variables. The best-fitting model for each subject area was identified using the Akaike Information Criterion (AIC), where a smaller value indicates a better fitting model (Akaike, 1974) and where a more complex model must have shown an AIC difference of twice the difference in the number of additional parameters in order for it to be selected, over a more parsimonious model.

Results

Descriptive Statistics

Sample sizes for college courses in each subject area ranged from relatively small for very specific subject areas, such as engineering ($n = 13,214$), to quite large for the more general subject areas, such as social science ($n = 115,324$; see Table 1). Engineering, mathematics, English, and the natural sciences had the greatest number of examinees taking at least one AP Exam within that subject. Across subject area, the percentage of students taking at least one AP Exam in the subject area of interest ranged from 6.2 percent for art and music to 68.4 percent for engineering.

Table 2 (see page 37) shows the composition of the nine subject area samples by gender, racial or ethnic identity, best spoken language, and highest parental education level. The subject areas with the greatest gender differences were engineering (80.8 percent male), computer science (63.4 percent male), and world language (63.1 percent female). Asian/Pacific Islander students made up a greater proportion of the engineering sample and a smaller proportion of the history sample than they did other subject areas. In addition, Black/African American students made up a relatively smaller proportion of the engineering sample than they did other subject areas. Differences in the composition of the subject area samples in terms of best spoken language were quite small, even for the world language sample. Finally, the engineering and world language samples both had relatively fewer students whose parents' highest educational level was a high school diploma or less and relatively more students with at least one parent having earned a graduate degree, relative to the other subject area samples.

Table 1**AP[®] Examinees by Subject Area Sample**

Sample	AP Exams Taken	<i>n</i>	%
Mathematics	Total	101,120	100.0
	No mathematics	67,539	66.8
	At least one mathematics	33,581	33.2
	Calculus AB	22,658	22.4
	Calculus BC	7,711	7.6
	Statistics	9,200	9.1
Computer Science	Total	18,253	100.0
	No computer science	17,111	93.7
	At least one computer science	1,142	6.3
	Computer Science A	835	4.6
	Computer Science AB	469	2.6
Engineering	Total	13,214	100.0
	No math or natural science	4,174	31.6
	At least one math	8,009	60.6
	At least one natural science	6,186	46.8
	At least one math or natural science	9,040	68.4
Natural Science	Total	91,596	100.0
	No natural science	64,457	70.4
	At least one natural science	27,139	29.6
	Biology	12,698	13.9
	Chemistry	10,138	11.1
	Environmental Science	2,960	3.2
	Physics B	6,174	6.7
	Physics C: Electricity & Magnetism	1,704	1.9
	Physics C: Mechanics	3,745	4.1
Social Science	Total	115,324	125.9
	No social science	88,383	76.6
	At least one social science	26,941	23.4
	Macroeconomics	7,191	6.2
	Microeconomics	3,863	3.3
	Comparative Government	1,423	1.2
	U.S. Government	16,109	14.0
	Human Geography	1,306	1.1
	Psychology	8,809	7.6
History	Total	48,729	100.0
	No history	34,701	71.2
	At least one history	14,028	28.8
	U.S. History	12,092	24.8
	European History	3,353	6.9
	World History	2,606	5.3
English	Total	102,375	100.0
	No English	70,391	68.8
	At least one English	31,984	31.2
	English Language & Composition	18,868	18.4
	English Literature & Composition	24,692	24.1

Table 1 (continued)AP[®] Examinees by Subject Area Sample

Sample	AP Exams Taken	<i>n</i>	%
World Language	Total	39,618	100.0
	No world language	32,189	81.2
	At least one world language	7,429	18.8
	French Language	1,695	4.3
	French Literature	209	0.5
	German Language	305	0.8
	Italian Language & Culture	100	0.3
	Spanish Language	4,714	11.9
	Spanish Literature	550	1.4
	Latin Literature	347	0.9
	Latin Vergil	486	1.2
	Art & Music	Total	52,677
No art or music		49,430	93.8
At least one art or music		3,247	6.2
Art History		1,016	1.9
Studio Drawing		792	1.5
Studio Art: 2-D Design		691	1.3
Studio Art: 3-D Design		97	0.2
Music Theory		931	1.8

Note: Subject area sample consists of all students taking college courses in the relevant subject area.

The percentages of students for each subject area by high school affiliation, urbanicity, estimated number of AP courses offered, and the mean numbers of students enrolled per grade are shown in Table 3 (see page 38). Public high school students tended to comprise a relatively smaller proportion of the world language sample than they did other subject areas, while private school (both Catholic and non-Catholic) students tended to make up relatively more of the world language sample than they did other subject areas. Students from suburban high schools tended to make up a greater percentage of the computer science sample than other subject areas and urban high school students made up a relatively smaller proportion of the computer science sample than other subject areas. The engineering sample tended to be comprised of relatively more students from high schools offering 21 or more AP courses and relatively fewer students from high schools with between 1 and 5 AP courses, compared to other subject areas; the English sample contained fewer students from high schools with 21 or more AP courses offered, relative to other subject areas. Across all subject areas, larger high schools (as measured by mean enrollment per grade) tended to account for greater proportions of students.

The college characteristics considered in this study were control, admittance rate, and undergraduate enrollment (see Table 4, page 39). Almost all of the subject area samples were comprised of approximately 70 percent public college attendees, except for the world language sample, which was comprised of only 56.0 percent public college attendees. The percentages of students by admittance rate were fairly consistent across subject area samples, whereby most students attended schools that admitted between 50 percent and 75 percent of students. In this study, a greater proportion of world language students tended to come from colleges with low admittance rates (under 50 percent) and fewer tended to come from those with high admittance rates (75 percent or greater), relative to other subject areas. In addition, a smaller proportion of computer science students were from colleges with low admittance rates than other subject area samples. Finally, as with mean high school enrollment per grade, each

of the nine subject areas consisted of a greater proportion of students from larger colleges (undergraduate enrollment $\geq 7,500$) than from smaller ones (undergraduate enrollment $< 7,500$). Those attending large colleges (with 15,000 or more undergraduates) tended to make up more of the engineering and less of the computer science samples, relative to other subject areas.

Cross-Classified Multilevel Modeling

To establish support for the use of cross-classified multilevel models, null models (i.e., models without any predictors) with random effects for the high school and college attended were estimated. Thus, for each subject area the null model was used to partition variance in SGPA into a component between high schools, a component between colleges, and the remaining (residual) variance within high school-college combinations. The formulation of the models in this study assumed that the random interaction between high school and college attended was zero; in other words, the separate main effects of high school and college attended adequately accounted for random variation in SGPA.

Table 5 shows the covariance parameter estimates from the null model for each subject area.³ These parameters separately indicated how much random variation was accounted for uniquely by high school and college. The larger these values were, the more support there was for the use of cross-classified multilevel modeling. The percent of variance accounted for by high school attended ranged from 2.0 percent for art and music to 4.6 percent for natural science. The college attended always accounted for substantially more variance than high school attended, ranging from 7.2 percent for mathematics to 12.1 percent for history. This result supported the hypothesis that SGPA varies more by college attended than by high school attended — presumably because SGPA was observed at the college level. The main effects for high school and college random effects accounted for SGPA variance between 10.4 percent for art and music and 15.3 percent for history, which supported the appropriateness of the cross-classification of the random effects.

Table 5
Cross-Classified Multilevel Model Covariance Parameter Estimates for Null Models

Subject	n	Intercept				Residual	
		High School		College		Est.	PVAF
		Est.	PVAF	Est.	PVAF		
Mathematics	101,120	0.0441	3.9	0.0812	7.2	1.0065	88.9
Computer Science	18,253	0.0339	3.0	0.1116	9.8	0.9967	87.3
Engineering	13,214	0.0232	3.1	0.0696	9.3	0.6554	87.6
Natural Science	91,596	0.0420	4.6	0.0781	8.5	0.8013	87.0
Social Science	115,324	0.0298	3.7	0.0679	8.4	0.7074	87.9
History	48,729	0.0309	3.2	0.1180	12.1	0.8242	84.7
English	102,375	0.0202	2.9	0.0763	10.9	0.6005	86.2
World Language	39,618	0.0229	3.6	0.0568	8.9	0.5594	87.5
Art & Music	52,677	0.0163	2.0	0.0670	8.4	0.7153	89.6

Note. PVAF = Percent of variance accounted for by the relevant covariance parameter. Null model refers to a model with no fixed effects specified, other than a common intercept, with random intercept effects for each high school and college under full-information maximum likelihood. All covariance parameter estimates were statistically significant at $\alpha = .01$ for each random effect in each subject area.

Mathematics Course Performance

Table 6 presents summary statistics for the 101,120 students who took at least one mathematics course in college. As self-reported HSGPA increased from the “C” range, mathematics SGPA, SAT-CR, SAT-M, SAT-W, and the number of mathematics AP Exams taken increased. Similarly, as the mean AP Exam grade in mathematics increased from 1 to 5 — recall that mean AP Exam grades in a subject area were computed and then rounded to the nearest score on the 1-to-5 scale — these variables also increased. In addition, all of the AP students, even those who received a mean AP Exam grade of 1, outperformed non-AP examinees on all measures.

Table 6

Means (Standard Deviations) on Academic Characteristics of Mathematics Sample

Academic Characteristic	<i>n</i>	Math SGPA	FYGPA	HSGPA	SAT-CR	SAT-M	SAT-W	AP Math Exams Taken
Total	101,120	2.72 (1.06)	2.99 (0.68)	3.61 (0.50)	558 (93.5)	584 (95.4)	552 (92.0)	0.39 (0.60)
High School GPA								
A- to A+	63,442	2.95 (0.97)	3.16 (0.60)	3.93 (0.24)	580 (90.6)	609 (89.5)	575 (88.9)	0.52 (0.65)
B- to B+	35,353	2.36 (1.09)	2.71 (0.68)	3.13 (0.23)	524 (86.4)	547 (90.2)	516 (83.1)	0.19 (0.45)
C- to C+	2,303	1.97 (1.15)	2.35 (0.76)	2.20 (0.19)	484 (82.9)	495 (86.4)	468 (79.5)	0.05 (0.24)
D or F	22	2.18 (1.03)	2.46 (0.78)	0.95 (0.55)	502 (121.5)	525 (118.1)	475 (121.3)	0.23 (0.53)
Mean AP Mathematics Exam Grade								
5	8,518	3.38 (0.74)	3.48 (0.48)	3.95 (0.34)	653 (79.8)	715 (54.4)	646 (78.3)	1.30 (0.51)
4	7,160	3.12 (0.86)	3.28 (0.54)	3.86 (0.37)	620 (77.8)	674 (57.6)	614 (78.1)	1.23 (0.46)
3	6,948	3.02 (0.89)	3.16 (0.56)	3.80 (0.39)	598 (79.9)	648 (61.1)	593 (80.0)	1.15 (0.38)
2	5,094	2.91 (0.94)	3.05 (0.61)	3.77 (0.41)	575 (80.5)	618 (63.9)	571 (78.3)	1.11 (0.32)
1	5,861	2.63 (1.04)	2.88 (0.66)	3.76 (0.42)	544 (81.8)	582 (71.5)	540 (79.7)	1.04 (0.19)
N ^a	67,539	2.56 (1.09)	2.88 (0.69)	3.50 (0.50)	535 (87.4)	550 (84.1)	528 (85.3)	0.00 (0.00)

^a Students took zero AP Exams in the subject area.

The results of the estimation of several cross-classified multilevel models for mathematics SGPA are presented in Table 7. Recall that Table 5 showed that there was significant ($p < .01$) random variation in the mathematics SGPA model across the 110 colleges and across the 7,480 high schools. According to the AIC, the best-fitting model was Model 3, which included gender, racial or ethnic identity, highest parental education level, HSGPA, SAT-CR, SAT-M, SAT-W, and the mean AP Exam grade in mathematics. The addition of the number of AP mathematics exams, as seen in Model 4, did not substantially improve the model fit of Model 3.

Table 7
Mathematics Cross-Classified Multilevel Model Results

Variable	Value / Group	Model 1	Model 2	Model 3	Model 4
		Est. (p)	Est. (p)	Est. (p)	Est. (p)
Fixed-Effects					
Intercept		2.659 (0.000)	2.616 (0.000)	2.597 (0.000)	2.597 (0.000)
Gender	Female	0.230 (0.000)	0.232 (0.000)	0.234 (0.000)	0.234 (0.000)
Racial or Ethnic Identity	American Indian	-0.175 (0.000)	-0.172 (0.000)	-0.170 (0.000)	-0.170 (0.000)
	Asian	0.027 (0.014)	0.008 (0.468)	0.009 (0.389)	0.010 (0.352)
	Black	-0.145 (0.000)	-0.154 (0.000)	-0.154 (0.000)	-0.154 (0.000)
	Hispanic	-0.117 (0.000)	-0.123 (0.000)	-0.118 (0.000)	-0.118 (0.000)
	Other	-0.040 (0.026)	-0.044 (0.012)	-0.043 (0.016)	-0.042 (0.017)
Highest Parental Ed. Level	Missing	0.053 (0.353)	0.062 (0.272)	0.061 (0.285)	0.061 (0.284)
	H.S. Diploma or Less	-0.058 (0.000)	-0.059 (0.000)	-0.058 (0.000)	-0.058 (0.000)
	Associate Degree	-0.035 (0.007)	-0.035 (0.007)	-0.033 (0.009)	-0.033 (0.009)
	Graduate Degree	0.025 (0.001)	0.022 (0.003)	0.021 (0.005)	0.021 (0.005)
Missing	0.018 (0.135)	0.015 (0.216)	0.013 (0.288)	0.013 (0.286)	
High School GPA ^a		0.571 (0.000)	0.556 (0.000)	0.550 (0.000)	0.549 (0.000)
SAT Critical Reading ^{a,b}		-0.053 (0.000)	-0.051 (0.000)	-0.054 (0.000)	-0.055 (0.000)
SAT Mathematics ^{a,b}		0.267 (0.000)	0.236 (0.000)	0.217 (0.000)	0.217 (0.000)
SAT Writing ^{a,b}		0.080 (0.000)	0.078 (0.000)	0.075 (0.000)	0.075 (0.000)
Number of Math AP Exams			0.124 (0.000)		-0.017 (0.181)
Mean AP Math Exam Grade	1			-0.004 (0.756)	0.014 (0.468)
	2			0.153 (0.000)	0.172 (0.000)
	3			0.196 (0.000)	0.215 (0.000)
	4			0.211 (0.000)	0.232 (0.000)
	5			0.361 (0.000)	0.383 (0.000)
Random Parameters					
Intercept	College	0.070 (0.000)	0.070 (0.000)	0.070 (0.000)	0.070 (0.000)
Intercept	High School	0.028 (0.000)	0.028 (0.000)	0.028 (0.000)	0.028 (0.000)
Residual		0.859 (0.000)	0.859 (0.000)	0.859 (0.000)	0.859 (0.000)
AIC (model parameters)		274,401 (19)	273,991 (20)	273,488 (24)	273,488 (25)
<p>Note. The reference group was white males whose parents' highest education level was a bachelor's degree and who took zero subject area AP Exams. Models were estimated based on 101,120 students from 7,535 high schools attending 110 colleges.</p> <p>^a Variable was grand-mean centered. ^b Variable was divided by 100.</p>					

The chosen model (Model 3) showed that on average and controlling for the other variables in the model, females earned expected mathematics SGPA of 0.23 higher than males; American Indian, Black/African American, and Hispanic/Latino students underperformed white students by between 0.12 and 0.17 points on expected mathematics SGPA. Students whose highest parental education level was a high school diploma or less or an associate degree earned expected mathematics SGPA of 0.06 and 0.03 points, respectively, less than those whose parents' highest education level was a bachelor's degree (the reference level) and those with at least one parent earning a graduate degree outperformed the reference group by 0.02 on expected mathematics SGPA.

In terms of the academic predictors, as HSGPA increased one point, expected mathematics SGPA increased by 0.55. On average, students who earned SAT-M and SAT-W scores of 100 points greater than the population mean were expected to earn mathematics SGPA of 0.22 and 0.08 higher, respectively. In other words, the marginal effect of SAT-M on expected mathematics SGPA — controlling for the other variables in the model — is almost three times as large as the marginal effect of SAT-W on expected mathematics SGPA. This was unsurprising

given that the knowledge and skills measured in first-year college mathematics courses overlaps more with those measured by the SAT-M than by the SAT-W. There was a small, but significant negative effect ($p < .01$) for SAT-CR on expected mathematics SGPA, with a parameter estimate of -0.05 associated with a 100-point increase in SAT-CR.

While students with a mean AP mathematics exam grade of 1 did not significantly outperform non-AP examinees in expected mathematics SGPA ($p > .01$), those with higher mean AP Exam grades in mathematics significantly ($p < .01$) and substantially outperformed non-AP examinees in mathematics. Expected mathematics SGPA differences ranged from 0.15 to 0.36 for those with mean AP mathematics exam grades of 2 to 5, respectively.

Computer Science Course Performance

The sample of students taking at least one college course in computer science is described in Table 8. This was one of the smallest groups ($n = 18,253$) in this study, which also had one of the smallest subject area participation rates in the AP Exams of interest (6.4 percent). All variables presented in the table increased with an increase in HSGPA, but students in adjacent mean AP computer science exam grade categories did not generally differ significantly ($p < .01$) from each other. The small number of significant differences may be attributed to the relatively small sample size in each mean AP computer science exam category, which ranged from 88 to 394 examinees.

Table 8

Means (Standard Deviations) on Academic Characteristics of Computer Science Sample

Academic Characteristic	<i>n</i>	Comp. Sci. SGPA	FYGPA	HSGPA	SAT-CR	SAT-M	SAT-W	AP Comp. Sci. Exams Taken
Total	18,253	2.93 (1.06)	2.96 (0.69)	3.53 (0.52)	552 (95.5)	590 (100.9)	543 (93.5)	0.07 (0.29)
High School GPA								
A- to A+	10,229	3.14 (0.95)	3.18 (0.61)	3.92 (0.24)	581 (90.8)	624 (92.6)	574 (89.0)	0.09 (0.33)
B- to B+	7,391	2.68 (1.11)	2.70 (0.68)	3.11 (0.23)	518 (87.7)	550 (92.7)	508 (83.4)	0.05 (0.23)
C- to C+	629	2.37 (1.17)	2.34 (0.75)	2.20 (0.19)	477 (81.9)	494 (87.7)	458 (78.7)	0.01 (0.12)
D or F	4	n/r n/r	n/r n/r	n/r n/r	n/r n/r	n/r n/r	n/r n/r	n/r n/r
Mean AP Computer Science Exam Grade								
5	394	3.49 (0.76)	3.39 (0.63)	3.87 (0.36)	670 (73.7)	733 (52.4)	652 (79.0)	1.21 (0.41)
4	273	3.12 (0.94)	3.06 (0.71)	3.75 (0.42)	630 (77.0)	698 (59.2)	610 (81.6)	1.16 (0.36)
3	197	3.05 (0.97)	2.98 (0.65)	3.67 (0.42)	617 (72.2)	674 (64.8)	595 (79.4)	1.14 (0.35)
2	88	2.90 (1.01)	2.86 (0.75)	3.61 (0.46)	585 (91.6)	662 (62.2)	569 (79.1)	1.09 (0.29)
1	190	2.73 (1.10)	2.77 (0.66)	3.43 (0.49)	567 (83.1)	631 (74.5)	550 (85.2)	1.01 (0.07)
N ^a	17,111	2.91 (1.06)	2.95 (0.69)	3.52 (0.53)	547 (94.0)	583 (98.9)	539 (92.1)	0.00 0.00

n/r Not reported due to small sample size ($n < 15$).

^a Students took zero AP Exams in the subject area.

The cross-classified multilevel model results for computer science SGPA are presented in Table 9. The high school and college level accounted for significant variation ($p < .01$) on expected computer science SGPA (see Table 5) among the 3,488 high schools and the 106 colleges represented in this sample. As with mathematics, the model that best fit the data included all demographic and non-AP prior academic preparedness variables, along with mean AP Exam performance in computer science (Model 3). In addition, Model 3 showed that females outperformed males by 0.19 points on expected computer science SGPA; American Indian, Black/African American, and Hispanic/Latino students underperformed white students by between 0.13 and 0.45 points on expected computer science SGPA. The only significant difference ($p < .01$) on expected computer science SGPA for parental education level was for those whose parents earned a high school diploma or less; this group underperformed those whose parents' highest education level was a bachelor's degree by 0.06 in terms of expected computer science SGPA.

Table 9
Computer Science Cross-Classified Multilevel Model Results

Variable	Value / Group	Model 1	Model 2	Model 3	Model 4
		Est. (p)	Est. (p)	Est. (p)	Est. (p)
Fixed-Effects					
Intercept		2.975 (0.000)	2.965 (0.000)	2.963 (0.000)	2.963 (0.000)
Gender	Female	0.182 (0.000)	0.186 (0.000)	0.187 (0.000)	0.187 (0.000)
Racial or Ethnic Identity	American Indian	-0.453 (0.000)	-0.450 (0.000)	-0.449 (0.000)	-0.449 (0.000)
	Asian	-0.022 (0.403)	-0.026 (0.314)	-0.026 (0.323)	-0.025 (0.328)
	Black	-0.321 (0.000)	-0.321 (0.000)	-0.321 (0.000)	-0.321 (0.000)
	Hispanic	-0.134 (0.000)	-0.134 (0.000)	-0.133 (0.000)	-0.133 (0.000)
	Other	-0.032 (0.469)	-0.033 (0.444)	-0.032 (0.458)	-0.032 (0.457)
Highest Parental Ed. Level	Missing	-0.138 (0.290)	-0.133 (0.309)	-0.133 (0.310)	-0.133 (0.309)
	H.S. Diploma or Less	-0.058 (0.004)	-0.058 (0.004)	-0.058 (0.005)	-0.058 (0.005)
	Associate Degree	-0.029 (0.334)	-0.029 (0.344)	-0.028 (0.362)	-0.028 (0.367)
	Graduate Degree	0.025 (0.167)	0.025 (0.165)	0.024 (0.177)	0.024 (0.177)
Missing	-0.010 (0.733)	-0.012 (0.686)	-0.012 (0.680)	-0.012 (0.682)	
High School GPA ^a		0.485 (0.000)	0.487 (0.000)	0.486 (0.000)	0.486 (0.000)
SAT Critical Reading ^{a,b}		0.012 (0.340)	0.011 (0.399)	0.010 (0.461)	0.010 (0.464)
SAT Mathematics ^{a,b}		0.132 (0.000)	0.127 (0.000)	0.126 (0.000)	0.126 (0.000)
SAT Writing ^{a,b}		0.062 (0.000)	0.062 (0.000)	0.061 (0.000)	0.061 (0.000)
Number of Comp. Sci. AP Exams			0.112 (0.000)		-0.053 (0.530)
Mean AP Comp. Sci. Exam Grade	1			-0.009 (0.899)	0.044 (0.689)
	2			0.088 (0.393)	0.145 (0.291)
	3			0.113 (0.104)	0.172 (0.143)
	4			0.104 (0.082)	0.164 (0.147)
	5			0.285 (0.000)	0.348 (0.002)
Random Parameters					
Intercept	College	0.098 (0.000)	0.098 (0.000)	0.098 (0.000)	0.098 (0.000)
Intercept	High School	0.024 (0.000)	0.024 (0.000)	0.024 (0.000)	0.024 (0.000)
Residual		0.890 (0.000)	0.890 (0.000)	0.890 (0.000)	0.890 (0.000)
AIC (model parameters)		50,382 (19)	50,367 (20)	50,359 (24)	50,361 (25)

Note. The reference group was white males whose parents' highest education level was a bachelor's degree and who took zero subject area AP Exams. Models were estimated based on 18,253 students from 3,488 high schools attending 106 colleges.

^a Variable was grand-mean centered. ^b Variable was divided by 100.

Among the academic predictors in the model, an increase in one point on HSGPA was related to an increase in expected computer science SGPA of 0.49. A 100-point increase in SAT mathematics and writing was associated with increases of 0.13 and 0.06 in expected computer science SGPA, respectively. Students with a mean computer science AP Exam grade of 5 were the only group who outperformed the reference group (i.e., non-computer science AP examinees), with expected computer science SGPA that were 0.29 higher.

Engineering Course Performance

The sample of students who took engineering courses had the smallest sample size ($n = 13,214$) and the greatest AP Exam participation rate (68.4 percent) of the nine subject areas examined. Table 10 shows that as HSGPA increased so too did each of the other academic predictors. Examinees who received a mean grade of 3 or greater significantly outperformed each adjacent lower grade category on all measures ($p < .01$; e.g., those with a mean score of 5 outperformed those with a mean score of 4), except on HSGPA for those whose mean grade was 3 and on the number of AP mathematics and natural science exams for those whose mean grade was 5. Performance on engineering SGPA and each SAT section among non-AP examinees was generally indistinguishable from AP examinees with a mean mathematics and science AP Exam grade of 1.

Table 10
Means (Standard Deviations) on Academic Characteristics of Engineering Sample

Academic Characteristic	<i>n</i>	Engineering SGPA	FYGPA	HSGPA	SAT-CR	SAT-M	SAT-W	AP Math & Sci. Exams Taken
Total	13,214	3.16 (0.86)	2.96 (0.69)	3.75 (0.45)	587 (87.9)	652 (79.2)	573 (87.9)	1.48 (1.42)
High School GPA								
A- to A+	9,796	3.28 (0.77)	3.10 (0.64)	3.96 (0.24)	599 (85.6)	664 (74.5)	587 (85.2)	1.69 (1.43)
B- to B+	3,275	2.83 (0.99)	2.57 (0.69)	3.17 (0.21)	554 (85.4)	619 (81.3)	536 (83.3)	0.93 (1.23)
C- to C+	142	2.45 (1.14)	2.21 (0.80)	2.22 (0.18)	522 (81.4)	569 (86.5)	491 (79.3)	0.30 (0.71)
D or F	1	n/r n/r	n/r n/r	n/r n/r	n/r n/r	n/r n/r	n/r n/r	n/r n/r
Mean AP Mathematics and Natural Science Exam Grade								
5	2,280	3.56 (0.59)	3.45 (0.54)	3.94 (0.34)	652 (78.0)	722 (53.8)	639 (76.7)	2.52 (1.34)
4	2,690	3.37 (0.68)	3.19 (0.55)	3.86 (0.38)	620 (72.3)	692 (55.1)	610 (74.1)	2.40 (1.21)
3	2,055	3.18 (0.75)	2.96 (0.59)	3.79 (0.40)	592 (75.4)	662 (56.3)	578 (74.9)	2.04 (1.11)
2	1,204	3.00 (0.86)	2.77 (0.61)	3.76 (0.40)	568 (72.3)	632 (59.8)	555 (74.0)	1.78 (0.93)
1	811	2.82 (0.97)	2.59 (0.67)	3.72 (0.44)	533 (77.8)	596 (65.0)	525 (71.6)	1.32 (0.65)
N ^a	4,174	2.90 (0.98)	2.65 (0.71)	3.55 (0.50)	544 (82.0)	599 (74.8)	527 (80.0)	0.00 0.00

n/r Not reported due to small sample size ($n < 15$).

^a Students took zero AP Exams in the subject area.

Table 11 shows the results of the cross-classified multilevel model analysis of engineering SGPA. According to the AIC, Model 4 was the best fit and included demographic predictors, non-AP academic predictors, and both the number of AP Exams taken and the mean AP Exam grades received in mathematics and natural science. The covariance parameters for the 2,823 high schools and the 65 colleges were significant ($p < .01$). Females outperformed males by a margin of 0.10 on expected engineering SGPA; Asian/Pacific Islander, Black/African American, and Hispanic/Latino students and those reporting their racial or ethnic identity as other earned lower expected SGPAs in engineering than white students by between 0.06 and 0.24 points; and students whose highest parental education level was a high school diploma or less were expected to earn engineering SGPAs of 0.09 points less than those whose parents' highest education level was a bachelor's degree. The non-AP academic predictors that were significantly related ($p < .01$) to expected engineering SGPA were HSGPA, SAT-M, and SAT-W, with parameter estimates of 0.45, 0.08, and 0.04, respectively.

Table 11**Engineering Cross-Classified Multilevel Model Results**

Variable	Value / Group	Model 1	Model 2	Model 3	Model 4
		Est. (p)	Est. (p)	Est. (p)	Est. (p)
Fixed-Effects					
Intercept		3.078 (0.000)	3.041 (0.000)	3.013 (0.000)	3.013 (0.000)
Gender	Female	0.092 (0.000)	0.093 (0.000)	0.094 (0.000)	0.096 (0.000)
Racial or Ethnic Identity	American Indian	-0.250 (0.052)	-0.226 (0.078)	-0.197 (0.122)	-0.193 (0.130)
	Asian	-0.047 (0.035)	-0.073 (0.001)	-0.054 (0.015)	-0.062 (0.005)
	Black	-0.237 (0.000)	-0.243 (0.000)	-0.235 (0.000)	-0.237 (0.000)
	Hispanic	-0.131 (0.000)	-0.138 (0.000)	-0.127 (0.000)	-0.127 (0.000)
	Other	-0.123 (0.006)	-0.133 (0.003)	-0.124 (0.005)	-0.127 (0.004)
	Missing	0.035 (0.823)	0.065 (0.674)	0.071 (0.644)	0.072 (0.641)
Highest Parental Ed. Level	H.S. Diploma or Less	-0.095 (0.000)	-0.095 (0.000)	-0.089 (0.000)	-0.089 (0.000)
	Associate Degree	0.006 (0.858)	0.006 (0.856)	0.018 (0.553)	0.018 (0.566)
	Graduate Degree	0.009 (0.587)	0.002 (0.880)	0.000 (0.982)	-0.001 (0.940)
	Missing	-0.002 (0.932)	-0.009 (0.754)	-0.013 (0.654)	-0.014 (0.620)
High School GPA ^a		0.488 (0.000)	0.472 (0.000)	0.454 (0.000)	0.453 (0.000)
SAT Critical Reading ^{a,b}		-0.011 (0.372)	-0.014 (0.252)	-0.025 (0.037)	-0.026 (0.033)
SAT Mathematics ^{a,b}		0.157 (0.000)	0.128 (0.000)	0.084 (0.000)	0.080 (0.000)
SAT Writing ^{a,b}		0.054 (0.000)	0.049 (0.000)	0.039 (0.002)	0.038 (0.002)
Number of Math & Nat. Sci. AP Exams			0.050 (0.000)		0.020 (0.009)
Mean AP Math & Nat. Sci. Exam Grade	1			-0.065 (0.030)	-0.091 (0.004)
	2			0.021 (0.415)	-0.010 (0.712)
	3			0.142 (0.000)	0.107 (0.000)
	4			0.244 (0.000)	0.206 (0.000)
	5			0.375 (0.000)	0.337 (0.000)
Random Parameters					
Intercept	College	0.044 (0.000)	0.044 (0.000)	0.044 (0.000)	0.044 (0.000)
Intercept	High School	0.020 (0.000)	0.020 (0.000)	0.020 (0.000)	0.020 (0.000)
Residual		0.576 (0.000)	0.576 (0.000)	0.576 (0.000)	0.576 (0.000)
AIC (model parameters)		30,777 (19)	30,710 (20)	30,508 (24)	30,503 (25)

Note. The reference group was white males whose parents' highest education level was a bachelor's degree and who took zero subject area AP Exams. Models were estimated based on 13,214 students from 2,823 high schools attending 65 colleges.

^a Variable was grand-mean centered. ^b Variable was divided by 100.

With each additional AP Exam taken in mathematics or natural science, students' engineering SGPA's increased by an average of 0.02 points. Those who had mean AP math and natural science exam grades of 3, 4, or 5 were expected to outperform the non-AP group by 0.11, 0.21, and 0.34 points, respectively, while those who scored a 1 on average tended to underperform comparable non-AP math and natural science students by 0.09.

Natural Science Course Performance

Among the students taking natural science courses in this study ($n = 91,596$), the results showed a positive relationship between high school GPA and mean AP natural science exam grade with all other academic variables (see Table 12). The cross-classified multilevel model results show that the best-fitting model was Model 4, which included all possible variables (see Table 13). In addition, both covariance parameters associated with the random intercept effect for the 7,267 high schools and 110 colleges were significant ($p < .01$).

Table 12

Means (Standard Deviations) on Academic Characteristics of Natural Science Sample

Academic Characteristic	<i>n</i>	Nat. Sci. SGPA	FYGPA	HSGPA	SAT-CR	SAT-M	SAT-W	AP Nat. Sci. Exams Taken
Total	91,596	2.76 (0.95)	3.00 (0.68)	3.66 (0.48)	566 (93.6)	591 (94.4)	560 (92.4)	0.41 (0.73)
High School GPA								
A- to A+	61,460	2.95 (0.86)	3.17 (0.60)	3.94 (0.24)	585 (90.7)	612 (88.4)	580 (89.2)	0.51 (0.80)
B- to B+	28,558	2.38 (0.99)	2.68 (0.69)	3.14 (0.22)	529 (86.8)	550 (90.6)	521 (84.0)	0.21 (0.52)
C- to C+	1,566	2.05 (1.09)	2.34 (0.77)	2.21 (0.19)	491 (85.9)	500 (87.6)	470 (81.9)	0.05 (0.25)
D or F	12	n/r n/r	n/r n/r	n/r n/r	n/r n/r	n/r n/r	n/r n/r	n/r n/r
Mean AP Natural Science Exam Grade								
5	4,670	3.45 (0.63)	3.53 (0.47)	3.97 (0.33)	679 (72.0)	711 (61.3)	668 (73.4)	1.66 (0.87)
4	6,532	3.18 (0.72)	3.32 (0.52)	3.88 (0.37)	640 (72.2)	677 (65.6)	631 (75.1)	1.51 (0.75)
3	6,905	2.98 (0.79)	3.17 (0.57)	3.82 (0.39)	609 (74.3)	648 (68.2)	601 (76.2)	1.33 (0.62)
2	5,260	2.79 (0.85)	3.03 (0.59)	3.77 (0.41)	573 (74.8)	610 (73.4)	570 (76.0)	1.24 (0.54)
1	3,772	2.54 (0.95)	2.84 (0.66)	3.72 (0.44)	537 (81.2)	574 (79.9)	536 (81.1)	1.10 (0.35)
N ^a	64,457	2.65 (0.97)	2.92 (0.70)	3.59 (0.49)	547 (89.4)	567 (88.6)	540 (87.9)	0.00 0.00

n/r Not reported due to small sample size ($n < 15$).

^a Students took zero AP Exams in the subject area.

The female students in the sample outperformed the males in terms of expected natural science SGPA by 0.09; all students who indicated their racial or ethnic group were expected to underperform white students, by a range of 0.06 to 0.27 points. Those whose parents' highest education level was either a high school diploma or less or an associate degree underperformed those for whom at least one parent earned at least a bachelor's degree on expected natural science SGPA, with parameter estimates of -0.10 and -0.07, respectively.

Students with at least one parent earning a graduate degree slightly, but significantly ($p < .01$), outperformed the reference group by 0.02 on expected natural science GPA. The parameter estimates for HSGPA and each SAT section were all significant ($p < .01$), and positive. With each additional natural science AP Exam taken, students' expected natural science SGPA increased by 0.02. Finally, AP examinees in natural science whose mean exam grade was a 1 earned expected natural science SGPA of 0.13 lower than comparable non-AP natural science examinees; and those whose mean exam grades were 3, 4, and 5, outperformed that same reference group by 0.04, 0.13, and 0.26, respectively.

Table 13
Natural Science Cross-Classified Multilevel Model Results

Variable	Value / Group	Model 1	Model 2	Model 3	Model 4
		Est. (p)	Est. (p)	Est. (p)	Est. (p)
Fixed-Effects					
Intercept		2.753 (0.000)	2.731 (0.000)	2.725 (0.000)	2.725 (0.000)
Gender	Female	0.079 (0.000)	0.082 (0.000)	0.084 (0.000)	0.085 (0.000)
Racial or Ethnic Identity	American Indian	-0.123 (0.002)	-0.121 (0.002)	-0.122 (0.002)	-0.121 (0.002)
	Asian	-0.039 (0.000)	-0.057 (0.000)	-0.053 (0.000)	-0.056 (0.000)
	Black	-0.266 (0.000)	-0.270 (0.000)	-0.268 (0.000)	-0.268 (0.000)
	Hispanic	-0.175 (0.000)	-0.179 (0.000)	-0.174 (0.000)	-0.174 (0.000)
	Other	-0.084 (0.000)	-0.091 (0.000)	-0.089 (0.000)	-0.089 (0.000)
	Missing	-0.008 (0.885)	-0.003 (0.955)	-0.005 (0.931)	-0.006 (0.922)
Highest Parental Ed. Level	H.S. Diploma or Less	-0.102 (0.000)	-0.102 (0.000)	-0.102 (0.000)	-0.102 (0.000)
	Associate Degree	-0.072 (0.000)	-0.072 (0.000)	-0.071 (0.000)	-0.071 (0.000)
	Graduate Degree	0.025 (0.000)	0.022 (0.001)	0.020 (0.003)	0.020 (0.004)
	Missing	0.001 (0.918)	-0.002 (0.861)	-0.005 (0.673)	-0.005 (0.659)
High School GPA ^a		0.513 (0.000)	0.505 (0.000)	0.502 (0.000)	0.502 (0.000)
SAT Critical Reading ^{a, b}		0.072 (0.000)	0.069 (0.000)	0.060 (0.000)	0.060 (0.000)
SAT Mathematics ^{a, b}		0.157 (0.000)	0.146 (0.000)	0.139 (0.000)	0.138 (0.000)
SAT Writing ^{a, b}		0.080 (0.000)	0.079 (0.000)	0.076 (0.000)	0.076 (0.000)
Number of Nat. Sci. AP Exams			0.059 (0.000)		0.024 (0.002)
Mean AP Nat. Sci. Exam Grade	1			-0.102 (0.000)	-0.128 (0.000)
	2			0.002 (0.877)	-0.027 (0.075)
	3			0.074 (0.000)	0.043 (0.004)
	4			0.167 (0.000)	0.132 (0.000)
	5			0.301 (0.000)	0.263 (0.000)
Random Parameters					
Intercept	College	0.050 (0.000)	0.050 (0.000)	0.050 (0.000)	0.050 (0.000)
Intercept	High School	0.019 (0.000)	0.019 (0.000)	0.019 (0.000)	0.019 (0.000)
Residual		0.681 (0.000)	0.681 (0.000)	0.681 (0.000)	0.681 (0.000)
AIC (model parameters)		227,087 (19)	226,909 (20)	226,457 (24)	226,449 (25)

Note. The reference group was white males whose parents' highest education level was a bachelor's degree and who took zero subject area AP Exams. Models were estimated based on 91,596 students from 7,267 high schools attending 110 colleges.

^a Variable was grand-mean centered. ^b Variable was divided by 100.

Social Science Course Performance

The largest sample ($n = 115,324$) was the group who took at least one course in social science in their first year of college. As Table 14 shows, there was a positive relationship between all of the academic variables and HSGPA as it increased from C- to C+ to A- to A+ and as mean social science AP Exam grade increased from 2 to 5; AP examinees who received a mean grade of 1 did not significantly ($p < .01$) outperform the non-examinee group.

Table 14

Means (Standard Deviations) on Academic Characteristics of Social Science Sample

Academic Characteristic	<i>n</i>	Soc. Sci. SGPA	FYGPA	HSGPA	SAT-CR	SAT-M	SAT-W	AP Soc. Sci. Exams Taken
Total	115,324	2.94 (0.89)	2.99 (0.68)	3.59 (0.50)	559 (95.1)	575 (95.5)	553 (93.8)	0.34 (0.70)
High School GPA								
A- to A+	70,171	3.16 (0.79)	3.18 (0.60)	3.92 (0.24)	582 (92.4)	601 (89.9)	578 (90.7)	0.42 (0.76)
B- to B+	42,312	2.63 (0.93)	2.71 (0.69)	3.13 (0.23)	525 (87.1)	538 (89.1)	518 (84.2)	0.21 (0.56)
C- to C+	2,814	2.25 (1.04)	2.35 (0.77)	2.21 (0.19)	485 (84.5)	488 (85.6)	469 (80.3)	0.08 (0.34)
D or F	27	2.23 (1.02)	2.31 (0.81)	1.05 (0.47)	473 (134.1)	490 (105.8)	449 (122.8)	0.07 (0.27)
Mean AP Social Science Exam Grade								
5	4,006	3.58 (0.53)	3.55 (0.44)	3.94 (0.34)	685 (70.2)	684 (70.2)	672 (71.9)	1.49 (0.74)
4	7,117	3.36 (0.64)	3.34 (0.49)	3.82 (0.39)	641 (71.5)	650 (74.1)	629 (73.5)	1.51 (0.75)
3	7,319	3.17 (0.72)	3.17 (0.56)	3.74 (0.41)	607 (70.7)	618 (77.4)	598 (75.8)	1.44 (0.70)
2	5,718	2.95 (0.80)	2.98 (0.62)	3.67 (0.44)	566 (70.7)	583 (77.5)	564 (73.0)	1.38 (0.63)
1	2,781	2.65 (0.91)	2.73 (0.69)	3.54 (0.47)	509 (75.9)	533 (86.1)	514 (78.6)	1.28 (0.57)
N ^a	88,383	2.87 (0.91)	2.93 (0.70)	3.54 (0.51)	544 (92.1)	561 (93.4)	539 (90.8)	0.00 (0.00)

^a Students took zero AP Exams in the subject area.

The cross-classified multilevel model that fits the social science SGPA data the best was Model 4, which included all variables of interest (see Table 15). There was sufficient variation across the 7,857 high schools and the 110 colleges that the covariance term for the random intercept effect was significant ($p < .01$) for each level. Females in this sample, on average, earned social science SGPA's of 0.15 higher than males; American Indian, Asian/Pacific Islander, Black/African American, and Hispanic/Latino students were expected to underperform comparable white students by between 0.03 and 0.12; and those whose parents' highest level of education was either a high school diploma or less or an associate degree underperformed comparable children of parents who earned a bachelor's degree, while students with at least one parent having earned a graduate degree slightly outperformed that same reference group on expected social science SGPA. The HSGPA and SAT critical reading, mathematics, and writing section scores were significantly ($p < .01$) and positively related to expected social science SGPA, with parameter estimates of 0.49, 0.11, 0.07, and 0.07, respectively.

Table 15
Social Science Cross-Classified Multilevel Model Results

Variable	Value / Group	Model 1	Model 2	Model 3	Model 4
		Est. (p)	Est. (p)	Est. (p)	Est. (p)
Fixed-Effects					
Intercept		2.943 (0.000)	2.928 (0.000)	2.921 (0.000)	2.921 (0.000)
Gender	Female	0.150 (0.000)	0.151 (0.000)	0.152 (0.000)	0.152 (0.000)
Racial or Ethnic Identity	American Indian	-0.091 (0.007)	-0.090 (0.007)	-0.090 (0.007)	-0.090 (0.007)
	Asian	-0.021 (0.015)	-0.029 (0.001)	-0.026 (0.003)	-0.026 (0.003)
	Black	-0.116 (0.000)	-0.117 (0.000)	-0.115 (0.000)	-0.115 (0.000)
	Hispanic	-0.107 (0.000)	-0.111 (0.000)	-0.106 (0.000)	-0.106 (0.000)
	Other	-0.015 (0.270)	-0.017 (0.220)	-0.015 (0.269)	-0.015 (0.265)
	Missing	-0.041 (0.337)	-0.035 (0.407)	-0.036 (0.399)	-0.036 (0.396)
Highest Parental Ed. Level	H.S. Diploma or Less	-0.074 (0.000)	-0.074 (0.000)	-0.074 (0.000)	-0.074 (0.000)
	Associate Degree	-0.056 (0.000)	-0.055 (0.000)	-0.055 (0.000)	-0.055 (0.000)
	Graduate Degree	0.027 (0.000)	0.026 (0.000)	0.025 (0.000)	0.025 (0.000)
	Missing	0.001 (0.881)	0.001 (0.912)	0.002 (0.868)	0.001 (0.884)
High School GPA ^a		0.497 (0.000)	0.495 (0.000)	0.491 (0.000)	0.491 (0.000)
SAT Critical Reading ^{a, b}		0.125 (0.000)	0.120 (0.000)	0.113 (0.000)	0.113 (0.000)
SAT Mathematics ^{a, b}		0.072 (0.000)	0.071 (0.000)	0.068 (0.000)	0.068 (0.000)
SAT Writing ^{a, b}		0.076 (0.000)	0.073 (0.000)	0.072 (0.000)	0.072 (0.000)
Number of Soc. Sci. AP Exams			0.055 (0.000)		0.016 (0.027)
Mean AP Soc. Sci. Exam Grade	1			-0.064 (0.000)	-0.084 (0.000)
	2			0.024 (0.030)	0.003 (0.856)
	3			0.095 (0.000)	0.073 (0.000)
	4			0.159 (0.000)	0.136 (0.000)
	5			0.222 (0.000)	0.200 (0.000)
Random Parameters					
Intercept	College	0.039 (0.000)	0.039 (0.000)	0.039 (0.000)	0.039 (0.000)
Intercept	High School	0.015 (0.000)	0.015 (0.000)	0.015 (0.000)	0.015 (0.000)
Residual		0.597 (0.000)	0.597 (0.000)	0.597 (0.000)	0.597 (0.000)
AIC (model parameters)		270,297 (19)	270,088 (20)	269,812 (24)	269,810 (25)

Note. The reference group was white males whose parents' highest education level was a bachelor's degree and who took zero subject area AP Exams. Models were estimated based on 115,324 students from 7,857 high schools attending 110 colleges.

^a Variable was grand-mean centered. ^b Variable was divided by 100.

This was the only model that, according to the AIC, was best fit with parameters for the number of AP Exams in the relevant subject area and the mean performance on those exams, but whose parameter estimate for the number of exams was only marginally significant ($p = .027$). Students whose mean AP social science exam grade was a 1 significantly ($p < .01$) underperformed students who did not take any social science exams by a margin of 0.08 and those with mean exam grades of 3, 4, and 5 outperformed non-examinees by 0.07, 0.14, and 0.20, respectively, in terms of expected social science SGPA.

History Course Performance

There were 48,729 students in this study who took at least one college course in history. History students showed a positive relationship of HSGPA over the range of C- to C+ to A- to A+ with all of the academic predictors, and a similar positive relationship existed between those predictors and mean AP Exam grades in history over the range of 2 through 5 (see Table 16). Students whose mean AP Exam grade in history was a 1 did not outperform students who did not take any AP Exams in history in terms of SGPA. However, AP examinees who received a grade of 2, 3, 4, or 5 outperformed non-AP examinees in history SGPA.

Table 16
Means (Standard Deviations) on Academic Characteristics of History Sample

Academic Characteristic	<i>n</i>	History SGPA	FYGPA	HSGPA	SAT-CR	SAT-M	SAT-W	AP History Exams Taken
Total	48,729	2.82 (0.98)	2.92 (0.71)	3.54 (0.50)	550 (94.2)	562 (93.4)	542 (92.2)	0.37 (0.64)
High School GPA								
A- to A+	27,811	3.05 (0.88)	3.13 (0.63)	3.91 (0.23)	574 (92.6)	589 (88.7)	568 (90.5)	0.49 (0.70)
B- to B+	19,506	2.54 (1.01)	2.67 (0.70)	3.12 (0.23)	520 (85.9)	530 (86.9)	511 (82.2)	0.22 (0.51)
C- to C+	1,397	2.14 (1.08)	2.33 (0.76)	2.21 (0.18)	485 (84.9)	483 (82.2)	470 (80.5)	0.08 (0.32)
D or F	15	2.60 (1.32)	2.46 (0.86)	1.06 (0.46)	505 (133.9)	519 (107.9)	490 (107.8)	0.20 (0.56)
Mean AP History Exam Grade								
5	1,941	3.66 (0.47)	3.58 (0.39)	3.95 (0.35)	707 (64.1)	680 (71.5)	687 (69.5)	1.50 (0.55)
4	2,593	3.38 (0.63)	3.33 (0.50)	3.81 (0.39)	656 (66.3)	637 (76.7)	637 (71.8)	1.40 (0.55)
3	3,711	3.18 (0.74)	3.15 (0.58)	3.75 (0.40)	612 (66.0)	614 (76.2)	600 (71.1)	1.31 (0.50)
2	3,688	2.93 (0.87)	2.98 (0.64)	3.69 (0.43)	570 (66.5)	580 (77.4)	562 (69.4)	1.18 (0.39)
1	2,095	2.48 (1.03)	2.67 (0.71)	3.60 (0.45)	509 (70.5)	533 (83.1)	508 (72.4)	1.09 (0.29)
N ^a	34,701	2.70 (1.00)	2.84 (0.72)	3.46 (0.51)	527 (85.9)	544 (89.2)	521 (84.5)	0.00 (0.00)

^a Students took zero AP Exams in the subject area.

Model 3 was the best fit for the history SGPA data (see Table 17). The covariance parameters for the intercept were significant ($p < .01$) for both the high school and college levels, which showed that mean achievement in college history SGPA varied across the 5,691 high schools and 109 colleges represented by the sample. Female students tended to outperform male students by an average of 0.11 points on the expected history SGPA scale; American Indian, Black/African American, and Hispanic/Latino students were expected to underperform comparable white students by 0.17, 0.13, and 0.12 points, respectively. Students whose highest parental education level was a high school diploma or less or an associate degree tended to underperform the reference group of students whose parents earned a bachelor's degree by 0.11 and 0.08, respectively, while students of parents who earned a graduate or professional degree outperformed the reference group by 0.03 on expected history SGPA.

Table 17
History Cross-Classified Multilevel Model Results

Variable	Value / Group	Model 1	Model 2	Model 3	Model 4
		Est. (p)	Est. (p)	Est. (p)	Est. (p)
Fixed-Effects					
Intercept		2.896 (0.000)	2.862 (0.000)	2.844 (0.000)	2.844 (0.000)
Gender	Female	0.108 (0.000)	0.109 (0.000)	0.114 (0.000)	0.114 (0.000)
Racial or Ethnic Identity	American Indian	-0.174 (0.001)	-0.171 (0.002)	-0.170 (0.002)	-0.170 (0.002)
	Asian	0.043 (0.009)	0.034 (0.039)	0.032 (0.053)	0.032 (0.053)
	Black	-0.132 (0.000)	-0.134 (0.000)	-0.132 (0.000)	-0.132 (0.000)
	Hispanic	-0.125 (0.000)	-0.131 (0.000)	-0.124 (0.000)	-0.124 (0.000)
	Other	0.005 (0.822)	0.001 (0.974)	0.002 (0.934)	0.002 (0.937)
	Missing	-0.116 (0.104)	-0.102 (0.151)	-0.102 (0.152)	-0.102 (0.152)
Highest Parental Ed. Level	H.S. Diploma or Less	-0.109 (0.000)	-0.109 (0.000)	-0.109 (0.000)	-0.109 (0.000)
	Associate Degree	-0.077 (0.000)	-0.077 (0.000)	-0.076 (0.000)	-0.076 (0.000)
	Graduate Degree	0.036 (0.000)	0.034 (0.001)	0.032 (0.001)	0.032 (0.001)
	Missing	-0.015 (0.378)	-0.015 (0.355)	-0.016 (0.343)	-0.016 (0.343)
High School GPA ^a		0.483 (0.000)	0.476 (0.000)	0.471 (0.000)	0.471 (0.000)
SAT Critical Reading ^{a, b}		0.142 (0.000)	0.131 (0.000)	0.117 (0.000)	0.117 (0.000)
SAT Mathematics ^{a, b}		0.019 (0.003)	0.019 (0.003)	0.019 (0.003)	0.019 (0.003)
SAT Writing ^{a, b}		0.107 (0.000)	0.102 (0.000)	0.097 (0.000)	0.097 (0.000)
Number of History AP Exams			0.083 (0.000)		0.005 (0.769)
Mean AP History Exam Grade	1			-0.072 (0.000)	-0.077 (0.003)
	2			0.095 (0.000)	0.089 (0.000)
	3			0.160 (0.000)	0.154 (0.000)
	4			0.205 (0.000)	0.199 (0.000)
	5			0.304 (0.000)	0.297 (0.000)
Random Parameters					
Intercept	College	0.053 (0.000)	0.053 (0.000)	0.053 (0.000)	0.053 (0.000)
Intercept	High School	0.017 (0.000)	0.017 (0.000)	0.017 (0.000)	0.017 (0.000)
Residual		0.717 (0.000)	0.717 (0.000)	0.717 (0.000)	0.717 (0.000)
AIC (model parameters)		123,418 (19)	123,281 (20)	123,121 (24)	123,123 (25)

Note. The reference group was white males whose parents' highest education level was a bachelor's degree and who took zero subject area AP Exams. Models were estimated based on 48,729 students from 5,691 high schools attending 109 colleges.

^a Variable was grand-mean centered. ^b Variable was divided by 100.

Parameter estimates for the academic variables in this model largely followed the expected patterns. As HSGPA increased by one point, history SGPA increased on average by 0.47 points; and as SAT-CR, SAT-M, and SAT-W increased by 100 points each, expected history SGPA was greater by 0.12, 0.02, and 0.10, respectively. AP examinees in history whose mean exam grade was 1 had an expected history SGPA of 0.07 less than comparable non-AP examinees, while those earning at least a 2 on their AP Exams in history tended to outperform comparable non-AP examinees by between 0.10 and 0.30.

English Course Performance

The English sample included 102,375 students who took at least one English course. In general, there was a positive relationship among each of the academic predictors and HSGPA across the range of C- to C+ to A- to A+ (see Table 18). Students who received a mean AP English exam grade of at least a 2 significantly ($p < .01$) outperformed the reference group of students who did not take either AP Exam in English on all measures.

Table 18
Means (Standard Deviations) on Academic Characteristics of English Sample

Academic Characteristic	<i>n</i>	English SGPA	FYGPA	HSGPA	SAT-CR	SAT-M	SAT-W	AP English Exams Taken
Total	102,375	3.14 (0.83)	2.97 (0.68)	3.55 (0.50)	552 (94.4)	569 (96.0)	546 (93.6)	0.43 (0.69)
High School GPA								
A- to A+	59,046	3.34 (0.70)	3.18 (0.59)	3.91 (0.23)	577 (92.0)	597 (90.7)	573 (91.0)	0.58 (0.75)
B- to B+	40,412	2.89 (0.89)	2.71 (0.68)	3.12 (0.23)	520 (86.1)	534 (89.0)	513 (83.7)	0.23 (0.53)
C- to C+	2,891	2.51 (1.02)	2.37 (0.74)	2.20 (0.19)	483 (84.0)	484 (84.4)	466 (80.3)	0.07 (0.29)
D or F	26	2.52 (1.01)	2.44 (0.84)	1.01 (0.52)	485 (116.6)	482 (104.3)	460 (106.9)	0.12 (0.33)
Mean AP English Exam Grade								
5	3,146	3.66 (0.47)	3.56 (0.41)	3.97 (0.32)	725 (54.0)	684 (70.2)	711 (56.6)	1.57 (0.50)
4	7,394	3.52 (0.57)	3.38 (0.48)	3.87 (0.37)	664 (57.3)	646 (74.4)	654 (61.2)	1.45 (0.50)
3	11,671	3.37 (0.66)	3.21 (0.55)	3.78 (0.40)	605 (58.8)	609 (78.9)	599 (62.6)	1.36 (0.48)
2	8,457	3.15 (0.80)	2.94 (0.64)	3.68 (0.43)	534 (63.3)	556 (83.0)	533 (66.1)	1.24 (0.43)
1	1,316	2.74 (0.95)	2.58 (0.73)	3.55 (0.47)	459 (71.2)	497 (88.5)	462 (68.3)	1.16 (0.37)
N ^a	70,391	3.04 (0.86)	2.88 (0.69)	3.44 (0.51)	528 (86.2)	552 (93.4)	522 (85.4)	0.00 (0.00)

^a Students took zero AP Exams in the subject area.

The cross-classified multilevel model results supported Model 3 as the best-fitting model, which included the mean AP Exam grade in English but not the number of AP Exams taken in English (see Table 19). With students from 7,480 high schools attending the 110 colleges, there was significant variation in English SGPA across both of these levels ($p < .01$, for both levels). Many of the same relationships between demographic variables and performance, in terms of expected English SGPA, were found in this subject area as with other subject areas. Females tended to outperform comparable males by 0.18; American Indian, Black/African American, Hispanic/Latino, those reporting their racial or ethnic identity as other, and those not reporting their racial or ethnic identity, were expected to underperform the reference group of white students by between 0.05 and 0.16 points. Those whose parents' highest level of education was a high school diploma or less and those whose highest parental education was an associate degree underperformed children of parents who earned a bachelor's degree by 0.07 and 0.04, respectively, in terms of expected English SGPA.

Table 19
English Cross-Classified Multilevel Model Results

Variable	Value / Group	Model 1	Model 2	Model 3	Model 4
		Est. (p)	Est. (p)	Est. (p)	Est. (p)
Fixed-Effects					
Intercept		3.080 (0.000)	3.062 (0.000)	3.055 (0.000)	3.055 (0.000)
Gender	Female	0.185 (0.000)	0.181 (0.000)	0.179 (0.000)	0.179 (0.000)
Racial or Ethnic Identity	American Indian	-0.133 (0.000)	-0.132 (0.000)	-0.132 (0.000)	-0.132 (0.000)
	Asian	-0.017 (0.052)	-0.022 (0.015)	-0.020 (0.022)	-0.020 (0.023)
	Black	-0.161 (0.000)	-0.165 (0.000)	-0.161 (0.000)	-0.161 (0.000)
	Hispanic	-0.136 (0.000)	-0.141 (0.000)	-0.136 (0.000)	-0.136 (0.000)
	Other	-0.045 (0.001)	-0.046 (0.001)	-0.045 (0.001)	-0.045 (0.001)
	Missing	-0.148 (0.000)	-0.142 (0.001)	-0.141 (0.001)	-0.141 (0.001)
Highest Parental Ed. Level	H.S. Diploma or Less	-0.068 (0.000)	-0.068 (0.000)	-0.067 (0.000)	-0.067 (0.000)
	Associate Degree	-0.036 (0.000)	-0.036 (0.000)	-0.036 (0.000)	-0.036 (0.000)
	Graduate Degree	0.014 (0.018)	0.013 (0.022)	0.013 (0.028)	0.013 (0.028)
	Missing	-0.013 (0.167)	-0.014 (0.146)	-0.014 (0.140)	-0.014 (0.140)
High School GPA ^a		0.394 (0.000)	0.388 (0.000)	0.387 (0.000)	0.387 (0.000)
SAT Critical Reading ^{a, b}		0.038 (0.000)	0.031 (0.000)	0.024 (0.000)	0.024 (0.000)
SAT Mathematics ^{a, b}		-0.001 (0.736)	0.000 (0.983)	0.000 (0.907)	0.000 (0.908)
SAT Writing ^{a, b}		0.134 (0.000)	0.129 (0.000)	0.124 (0.000)	0.124 (0.000)
Number of English AP Exams			0.045 (0.000)		-0.005 (0.603)
Mean AP English Exam Grade	1			-0.084 (0.000)	-0.079 (0.001)
	2			0.049 (0.000)	0.054 (0.000)
	3			0.082 (0.000)	0.088 (0.000)
	4			0.112 (0.000)	0.119 (0.000)
	5			0.143 (0.000)	0.150 (0.000)
Random Parameters					
Intercept	College	0.037 (0.000)	0.037 (0.000)	0.037 (0.000)	0.037 (0.000)
Intercept	High School	0.011 (0.000)	0.011 (0.000)	0.011 (0.000)	0.011 (0.000)
Residual		0.528 (0.000)	0.528 (0.000)	0.528 (0.000)	0.528 (0.000)
AIC (model parameters)		227,221 (19)	227,096 (20)	226,995 (24)	226,997 (25)

Note. The reference group was white males whose parents' highest education level was a bachelor's degree and who took zero subject area AP Exams. Models were estimated based on 102,375 students from 7,480 high schools attending 110 colleges.

^a Variable was grand-mean centered. ^b Variable was divided by 100.

The academic variables of HSGPA, SAT-CR, and SAT-W had significant ($p < .01$) positive relationships with expected English SGPA, with parameter estimates of 0.39, 0.02, and 0.12, respectively. The students who earned a mean grade of 1 on the AP English exams were expected to underperform the reference group of non-examinees by 0.08; while those who earned means of 2, 3, 4, and 5 outperformed the reference group by 0.05, 0.08, 0.11, and 0.14, respectively, on expected English SGPA.

World Language Course Performance

Among the 39,618 world language students in this sample, there was a strong positive relationship between HSGPA and all academic variables (see Table 20). In terms of mean AP world language exam grade, students with mean scores of 1 significantly ($p < .01$) outperformed non-AP world language examinees. Each mean AP Exam group significantly ($p < .01$) outperformed the lower-scoring group on all measures, except that those whose mean scores were 4 and 5 were indistinguishable on world language SGPA, HSGPA, and SAT-M.

Table 20

Means (Standard Deviations) on Academic Characteristics of World Language Sample

Academic Characteristic	<i>n</i>	World. Lang. SGPA	FYGPA	HSGPA	SAT-CR	SAT-M	SAT-W	AP World Lang. Exams Taken
Total	39,618	3.29 (0.80)	3.17 (0.61)	3.67 (0.47)	592 (96.7)	595 (93.9)	587 (95.1)	0.21 (0.47)
High School GPA								
A- to A+	26,888	3.46 (0.65)	3.33 (0.51)	3.94 (0.24)	614 (91.6)	618 (87.9)	610 (89.5)	0.26 (0.51)
B- to B+	12,145	2.93 (0.93)	2.84 (0.64)	3.14 (0.22)	548 (90.2)	550 (88.0)	541 (87.5)	0.11 (0.35)
C- to C+	576	2.57 (1.07)	2.49 (0.69)	2.22 (0.19)	501 (87.1)	502 (88.1)	483 (82.9)	0.02 (0.15)
D or F	9	n/r n/r	n/r n/r	n/r n/r	n/r n/r	n/r n/r	n/r n/r	n/r n/r
Mean AP World Language Exam Grade								
5	1,269	3.74 (0.42)	3.52 (0.43)	3.92 (0.34)	683 (95.7)	669 (89.3)	683 (93.0)	1.27 (0.52)
4	1,561	3.65 (0.49)	3.46 (0.44)	3.87 (0.36)	666 (87.9)	657 (83.4)	668 (82.3)	1.20 (0.42)
3	2,344	3.61 (0.51)	3.40 (0.45)	3.85 (0.36)	643 (80.7)	645 (80.4)	648 (77.1)	1.10 (0.31)
2	1,476	3.52 (0.59)	3.30 (0.49)	3.80 (0.39)	615 (74.8)	626 (77.8)	619 (74.5)	1.06 (0.27)
1	779	3.30 (0.74)	3.13 (0.56)	3.73 (0.39)	590 (82.9)	596 (85.2)	589 (80.3)	1.02 (0.14)
N ^a	32,189	3.22 (0.83)	3.12 (0.62)	3.63 (0.48)	580 (94.5)	584 (92.4)	574 (92.0)	0.00 0.00

n/r Not reported due to small sample size ($n < 15$).

^a Students took zero AP Exams in the subject area.

The cross-classified multilevel modeling results across 5,402 high schools and 105 colleges for world language SGPA are presented in Table 21. This is the only model in which best spoken language was added to the models to account for potential differences for native speakers. Students whose best spoken language was English alone served as the reference group, which was compared to students whose best spoken language was another language or a combination of English and another language. The best-fitting model was Model 3, which included every variable except for the number of AP Exams taken in the world languages. There was significant variation ($p < .01$) across the high schools and colleges to warrant including a random effect for each level. The results showed that females outperformed males by 0.18 on expected world language SGPA. American Indian and Black/African American students were expected to underperform comparable white students by 0.24 and 0.15,

respectively, while Asian/Pacific Islander students were expected to outperform that same reference group by 0.10 on world language SGPA. Those whose parents earned a high school diploma or less tended to underperform students whose highest parental educational attainment was a bachelor's degree by 0.04. In addition, students whose best spoken language was a combination of English and another language and students whose best spoken language was another language outperformed students whose best language was English by 0.13 and 0.35 in terms of expected world language SGPA, respectively.

Table 21
World Language Cross-Classified Multilevel Model Results

Variable	Value / Group	Model 1	Model 2	Model 3	Model 4
		Est. (p)	Est. (p)	Est. (p)	Est. (p)
Fixed-Effects					
Intercept		3.078 (0.000)	3.068 (0.000)	3.067 (0.000)	3.067 (0.000)
Gender	Female	0.192 (0.000)	0.184 (0.000)	0.183 (0.000)	0.183 (0.000)
Racial or Ethnic Identity	American Indian	-0.248 (0.000)	-0.245 (0.000)	-0.239 (0.000)	-0.239 (0.000)
	Asian	0.096 (0.000)	0.102 (0.000)	0.104 (0.000)	0.104 (0.000)
	Black	-0.147 (0.000)	-0.146 (0.000)	-0.145 (0.000)	-0.145 (0.000)
	Hispanic	0.037 (0.015)	0.014 (0.374)	0.008 (0.603)	0.008 (0.597)
	Other	-0.004 (0.843)	-0.007 (0.709)	-0.007 (0.705)	-0.007 (0.705)
	Missing	-0.037 (0.601)	-0.028 (0.687)	-0.029 (0.685)	-0.029 (0.685)
Highest Parental Ed. Level	H.S. Diploma or Less	-0.039 (0.000)	-0.039 (0.000)	-0.039 (0.000)	-0.039 (0.000)
	Associate Degree	-0.028 (0.105)	-0.025 (0.136)	-0.025 (0.135)	-0.025 (0.134)
	Graduate Degree	0.011 (0.216)	0.008 (0.367)	0.007 (0.413)	0.007 (0.413)
	Missing	0.020 (0.245)	0.015 (0.360)	0.014 (0.396)	0.014 (0.396)
Best Spoken Language	English and Another	0.156 (0.000)	0.135 (0.000)	0.130 (0.000)	0.130 (0.000)
	Another	0.370 (0.000)	0.349 (0.000)	0.346 (0.000)	0.347 (0.000)
	Missing	-0.017 (0.486)	-0.019 (0.434)	-0.018 (0.444)	-0.018 (0.445)
High School GPA ^a		0.456 (0.000)	0.454 (0.000)	0.453 (0.000)	0.452 (0.000)
SAT Critical Reading ^{a, b}		-0.026 (0.000)	-0.027 (0.000)	-0.028 (0.000)	-0.028 (0.000)
SAT Mathematics ^{a, b}		0.089 (0.000)	0.087 (0.000)	0.086 (0.000)	0.086 (0.000)
SAT Writing ^{a, b}		0.121 (0.000)	0.113 (0.000)	0.110 (0.000)	0.110 (0.000)
Number of World Lang. AP Exams			0.120 (0.000)		-0.010 (0.659)
Mean AP World Lang. Exam Grade	1			0.020 (0.442)	0.030 (0.386)
	2			0.141 (0.000)	0.152 (0.000)
	3			0.176 (0.000)	0.187 (0.000)
	4			0.171 (0.000)	0.183 (0.000)
	5			0.212 (0.000)	0.224 (0.000)
Random Parameters					
Intercept	College	0.035 (0.000)	0.035 (0.000)	0.035 (0.000)	0.035 (0.000)
Intercept	High School	0.018 (0.000)	0.018 (0.000)	0.018 (0.000)	0.018 (0.000)
Residual		0.480 (0.000)	0.480 (0.000)	0.480 (0.000)	0.480 (0.000)
AIC (model parameters)		84,878 (22)	84,672 (23)	84,614 (27)	84,616 (28)

Note. The reference group was white males whose best spoken language is English only, whose parents' highest education level was a bachelor's degree and who took zero subject area AP Exams. Models were estimated based on 39,618 students from 5,402 high schools attending 105 colleges.

^a Variable was grand-mean centered. ^b Variable was divided by 100.

The academic variables were all significantly related to expected world language SGPA ($p < .01$). Specifically, HSGPA (0.45), SAT-M (0.09), and SAT-W (0.11) had a positive relationship and SAT-CR (-0.03) had a small negative relationship with expected world language SGPA. For mean world language AP Exam grade, students with a grade of 2 through 5, on average, outperformed non-examinees by between 0.14 and 0.21, while students who received a mean AP Exam grade of 1 did not perform significantly differently from non-examinees ($p > .01$) on expected world language SGPA.

Art and Music Course Performance

In the art and music sample ($n = 52,677$), there was a clear, positive relationship among the academic variables and each level of HSGPA (see Table 22). Those with mean AP art and music exam grades of 3, 4, and 5, outperformed non-examinees in terms of SAT and HSGPA.

Table 22

Means (Standard Deviations) on Academic Characteristics of Art and Music Sample

Academic Characteristic	<i>n</i>	Art & Music SGPA	FYGPA	HSGPA	SAT-CR	SAT-M	SAT-W	AP Art & Mus. Exams Taken
Total	52,677	3.25 (0.89)	3.03 (0.67)	3.58 (0.50)	560 (96.6)	573 (95.6)	555 (94.7)	0.07 (0.27)
High School GPA								
A- to A+	31,641	3.46 (0.74)	3.24 (0.57)	3.92 (0.24)	585 (93.7)	599 (90.0)	581 (91.2)	0.08 (0.29)
B- to B+	19,674	2.98 (0.99)	2.75 (0.67)	3.12 (0.23)	526 (88.0)	536 (88.9)	519 (85.4)	0.05 (0.24)
C- to C+	1,353	2.49 (1.14)	2.37 (0.74)	2.21 (0.19)	483 (85.0)	483 (84.9)	468 (82.0)	0.02 (0.16)
D or F	9	n/r n/r	n/r n/r	n/r n/r	n/r n/r	n/r n/r	n/r n/r	n/r n/r
Mean AP Art and Music Exam Grade								
5	560	3.66 (0.51)	3.49 (0.49)	3.82 (0.41)	656 (86.7)	662 (80.0)	654 (85.4)	1.07 (0.26)
4	811	3.55 (0.56)	3.34 (0.52)	3.72 (0.43)	622 (84.6)	620 (81.2)	619 (82.2)	1.14 (0.38)
3	1,075	3.39 (0.64)	3.21 (0.55)	3.65 (0.44)	598 (87.9)	601 (85.4)	592 (87.0)	1.10 (0.34)
2	634	3.23 (0.75)	3.03 (0.63)	3.55 (0.46)	573 (90.0)	571 (90.5)	568 (88.0)	1.03 (0.18)
1	167	3.13 (0.95)	2.89 (0.71)	3.52 (0.52)	550 (87.2)	547 (90.9)	541 (91.1)	1.01 (0.08)
N ^a	49,430	3.24 (0.90)	3.02 (0.67)	3.57 (0.50)	557 (96.1)	570 (95.5)	552 (94.1)	0.00 0.00

n/r Not reported due to small sample size ($n < 15$).

^a Students took zero AP Exams in the subject area.

The sample of students who took courses in art and music in their first year at one of 109 colleges came from 5,995 high schools. The best-fitting model was Model 3, which showed significant variation ($p > .01$) across those colleges and high schools on art and music SGPA such that the inclusion of random effect terms for both high school and college was appropriate (see Table 23). There was a significant gender effect, whereby females were expected to outperform comparable males by 0.15 ($p > .01$) on art and music SGPA. American Indian, Black/African American, Hispanic/Latino students, and those who did not report

their racial or ethnic identity underperformed the white reference group by 0.15, 0.12, 0.08, and 0.18, respectively, in terms of expected art and music SGPA. Students whose highest parental education level was either a high school diploma or less or an associate degree tended to underperform students of parents who earned a bachelor's degree by 0.10 and 0.07, respectively, with students of parents who earned a graduate degree earning higher expected art and music SGPAs by 0.02.

Table 23
Art and Music Cross-Classified Multilevel Model Results

Variable	Value / Group	Model 1	Model 2	Model 3	Model 4
		Est. (p)	Est. (p)	Est. (p)	Est. (p)
Fixed-Effects					
Intercept		3.246 (0.000)	3.244 (0.000)	3.243 (0.000)	3.243 (0.000)
Gender	Female	0.147 (0.000)	0.146 (0.000)	0.146 (0.000)	0.146 (0.000)
Racial or Ethnic Identity	American Indian	-0.152 (0.005)	-0.151 (0.005)	-0.151 (0.005)	-0.151 (0.005)
	Asian	0.009 (0.504)	0.009 (0.547)	0.008 (0.558)	0.008 (0.551)
	Black	-0.123 (0.000)	-0.122 (0.000)	-0.122 (0.000)	-0.122 (0.000)
	Hispanic	-0.080 (0.000)	-0.081 (0.000)	-0.080 (0.000)	-0.080 (0.000)
	Other	-0.036 (0.088)	-0.038 (0.070)	-0.037 (0.072)	-0.037 (0.075)
	Missing	-0.185 (0.005)	-0.182 (0.005)	-0.183 (0.005)	-0.182 (0.005)
Highest Parental Ed. Level	H.S. Diploma or Less	-0.097 (0.000)	-0.097 (0.000)	-0.097 (0.000)	-0.097 (0.000)
	Associate Degree	-0.069 (0.000)	-0.069 (0.000)	-0.069 (0.000)	-0.069 (0.000)
	Graduate Degree	0.024 (0.007)	0.023 (0.009)	0.023 (0.010)	0.023 (0.010)
	Missing	-0.013 (0.399)	-0.013 (0.381)	-0.013 (0.392)	-0.013 (0.393)
High School GPA ^a		0.441 (0.000)	0.441 (0.000)	0.441 (0.000)	0.441 (0.000)
SAT Critical Reading ^{a, b}		0.027 (0.000)	0.026 (0.000)	0.026 (0.000)	0.026 (0.000)
SAT Mathematics ^{a, b}		0.047 (0.000)	0.047 (0.000)	0.046 (0.000)	0.046 (0.000)
SAT Writing ^{a, b}		0.079 (0.000)	0.079 (0.000)	0.078 (0.000)	0.078 (0.000)
Number of Art & Music AP Exams			0.055 (0.000)		-0.042 (0.368)
Mean AP Art & Music Exam Grade	1			-0.044 (0.482)	-0.002 (0.978)
	2			-0.012 (0.715)	0.031 (0.589)
	3			0.054 (0.031)	0.100 (0.078)
	4			0.125 (0.000)	0.172 (0.004)
	5			0.139 (0.000)	0.183 (0.002)
Random Parameters					
Intercept	College	0.031 (0.000)	0.031 (0.000)	0.031 (0.000)	0.031 (0.000)
Intercept	High School	0.009 (0.000)	0.009 (0.000)	0.009 (0.000)	0.009 (0.000)
Residual		0.643 (0.000)	0.643 (0.000)	0.643 (0.000)	0.643 (0.000)
AIC (model parameters)		127,207 (19)	127,192 (20)	127,179 (24)	127,181 (25)

Note. The reference group was white males whose parents' highest education level was a bachelor's degree and who took zero subject area AP Exams. Models were estimated based on 52,677 students from 5,995 high schools attending 109 colleges.

^a Variable was grand-mean centered. ^b Variable was divided by 100.

The academic variable with the largest marginal effect on expected art and music SGPA was HSGPA, which had a parameter estimate of 0.44. SAT-CR, SAT-M, and SAT-W all had smaller, but significant ($p < .01$) parameter estimates of 0.03, 0.05, and 0.08, respectively, for each 100-point increase in the relevant section score. Finally, students whose mean AP Exam grade in art and music was a 4 or 5 outperformed non-examinees in terms of expected art and music SGPA by 0.13 and 0.14, respectively.

Discussion

As has been evaluated by previous authors, students were hypothesized to perform better in same-discipline, first-year college courses with either greater mean AP Exam performance or more AP Exams taken in the discipline. Consistently across the nine subject areas, there was support for the claim that better mean performance on AP Exams was related to better performance in first-year courses in the discipline. In only a subset of the nine content areas was the number of exams in the discipline related to higher expected performance in the discipline.

When considering the cross-classified multilevel model results across the nine subject areas, a few patterns emerge clearly. For all nine subject areas, there was sufficient random variation in the corresponding SGPA to warrant including random effect terms for the intercept for both the high school and college attended. The intercept of each null model randomly varied across high schools and colleges, with college accounting for between two and four times the random variation in SGPA than high school for each of the nine content areas (see Table 5). This was expected and supports the notion that different grading practices and college course curricula accounted for more variation in grades than could be attributed to high school attended.

For all nine subject areas, there was sufficient random variation in the corresponding SGPA to warrant including random effect terms for the intercept for both the high school and college attended.

Each of the best-fitting models showed a significant relationship between the demographic and socioeconomic predictor variables and SGPA. Across each of the nine subject areas, females tended to outperform males on expected SGPA by between 0.08 and 0.23 points, beyond what would have been predicted by the other variables in the best-fitting models. Relative to white students and controlling for the other variables in the chosen models: American Indian students earned between 0.09 and 0.45 lower expected SGPAs in eight subject areas; Black/African American students earned between 0.12 and 0.32 lower expected SGPAs in all nine subject areas; and Hispanic/Latino students earned between 0.08 and 0.17 lower expected SGPAs in eight subject areas. Given the other variables in the best-fitting models, Asian students were expected to slightly outperform white students on world language SGPA — by 0.10 — but they were expected to slightly underperform white students on mathematics and natural and social science SGPA — by between 0.03 and 0.06. This general pattern of results is supported by Hargrove, Godin, and Dodd (2008) who examined seven AP Exams and found significant gender and racial or ethnic differences across AP and matched non-AP groups for overall first-year college GPA.

Students whose parents completed more schooling were expected to outperform in terms of SGPA those whose parents completed less schooling. Across all nine content areas, students whose parents completed a high school diploma or less were expected to earn SGPAs of between 0.11 and 0.04 lower than students whose parents' earned at most a bachelor degree (reference level) and who were otherwise comparable on the measures included in the chosen model. Those whose parents earned at most an associate degree were expected to underperform the reference group in six of

the nine content areas by margins ranging from 0.08 to 0.03. Students with at least one parent earning a complete graduate degree outperformed the reference group in five of the nine content areas, with expected differences in SGPA of 0.02 to 0.03. In terms of practical significance, these parameter estimates are all quite small, being at most one-third of the difference between a B (3.00) and a B+ (3.33) on the typical grade scale. The differences in parameter estimates across subject areas could be due to any number of phenomena, for example, differential degrees of self-selection into subject areas being related to parental education level.

Despite the significant effects that were found for the demographic and socioeconomic variables, the relationships with expected SGPA across subject areas were stronger for the academic variables. More specifically, holding constant the other variables in the model, HSGPA always had a strong, significant, positive relationship with each of the nine SGPA; and depending upon the subject area, at least two of the SAT sections had a significant positive relationship with SGPA. The SAT sections that were found to be significantly and positively related to each SGPA seemed to be those that were the closest in terms of content, such as SAT mathematics and mathematics SGPA.

The main relationship of interest in this study was the relationship among SGPA and participation in and performance on AP Exams in each subject area. For three of the nine subject areas (engineering, natural science, and social science), the best-fitting model included both the number of subject area AP Exams and the mean AP Exam grade for that subject area. The model that fit the remaining six subjects best only included the mean AP Exam grade for that subject area. In the two subject areas (engineering and natural science) where there was a significant, positive relationship between the number of AP Exams taken in that discipline and the corresponding SGPA, the effect was very small (0.02). Thus, it seems that, overall, mean AP performance was more predictive of SGPA than AP participation across the nine subject areas.

Of particular interest for professors in these disciplines and other college and university administrators may be the AP Exam grades that were associated with significantly greater SGPA. For seven of the nine subject areas, students with a mean AP Exam grade of 3 or better significantly outperformed the reference group of non-examinees in the relevant subject area. The two subject areas where students earning a 3 failed to outperform non-AP examinees — art and music and computer science — were also those with the smallest AP participation rates. In four of the nine content areas (mathematics, history, English, and world language), students whose mean AP Exam grade in the subject was a 2 significantly outperformed non-AP examinees in that discipline in terms of expected SGPA.

For seven of the nine subject areas, students with a mean AP Exam grade of 3 or better significantly outperformed the reference group of non-examinees in the relevant subject area.

Limitations and Directions for Future Research

As previously noted, this study was limited by the fact that identification of AP students was based on those who took the AP Exam, rather than the larger group of all students who were exposed to the AP curriculum by taking an AP course. The sub-group of students who took an AP course but not the corresponding AP Exam may behave differently in terms of their subject area performance (i.e., SGPA) and as such, inferences to that group cannot be made on the basis of these results. In the present study, students taking the AP course, but not the corresponding AP Exam, would have been included in the non-AP group. Future research should aim to make this distinction and ensure that the non-AP group is truly non-AP.

In addition, meaningful data on the number of years of high school course work students took in each subject area were difficult to disentangle from AP course work. Future studies would benefit from obtaining actual high school records and incorporating high school course-taking patterns — in both AP and other courses — into models such as those presented herein. In particular, more

detailed high school transcript data would enable researchers to disentangle the effect of AP from the effect of the number of years of study in a discipline as well as test for differences in the effects of AP course- and exam-taking from honors, advanced, and dual enrollment course work.

This study is also limited in that it relies on average course performance in subject areas (i.e., SGPA). This introduces a few challenges, chief among them: (a) each student's SGPA may be based on a different number of courses/credits; and (b) the courses within each subject area may vary in difficulty. The varying number of courses is likely to have led to a different degree of precision with which mean SGPA was estimated. Systematic differences in average course difficulty within a particular college and subject area should not have biased the results, as random effects associated with each college were included and each subject area was modeled separately.

Conclusions

The results of this study support the notion that AP Exam performance — above and beyond gender, racial or ethnic identity, socioeconomic status, and academic ability — was related to first-year college performance in each of the nine subject areas considered. The number of AP Exams taken by students tended not to improve the overall fit of the cross-classified multilevel model for subject area GPA (SGPA), after accounting for AP performance and the other variables included in the models. As the average score on AP Exams in each subject area increased, so did the expected SGPA tend

The results of this study support the notion that AP Exam performance — above and beyond gender, racial or ethnic identity, socioeconomic status, and academic ability — was related to first-year college performance in each of the nine subject areas considered.

to increase, relative to students who did not take an AP Exam. In other words, there was a positive effect of AP Exam performance across multiple domains. In particular, students whose mean AP Exam grade in a given subject area was at least a 3 outperformed non-AP examinees in seven out of nine subject areas. The consistent positive effect associated with mean scores on subject area AP Exams relative to students who took no exams across the subject areas considered serves as evidence that the Advanced Placement program seems to prepare students to do well in college-level course work.

References

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716–723. doi:10.1109/TAC.1974.1100705.
- Beretvas, S. N. (2008). Cross-classified random effects models. In O’Connell, A. A. and McCoach, D. B. (Eds.), *Multilevel modeling of educational data (quantitative methods in education and the behavioral science)*. Charlotte, NC: Information Age Publishing.
- Cassery, P. L. (1986). *Advanced Placement revisited* (College Board research report no. 86-6). New York, NY: The College Board.
- College Board, The. (2006). *2006 College-bound seniors: Total group profile report*. New York, NY: The College Board. Retrieved May 13, 2010, from http://www.collegeboard.com/prod_downloads/about/news_info/cbsenior/yr2006/national-report.pdf.
- College Board, The. (2007). *The College Board college handbook 2007* (44th ed.). New York, NY: The College Board.
- College Board, The. (2008). *The fourth annual AP® report to the nation*. New York, NY: The College Board. Retrieved May 13, 2010, from <http://professionals.collegeboard.com/profdownload/ap-report-to-the-nation-2008.pdf>.
- Dodd, B. G., Fitzpatrick, S. J., De Ayala, R. J., & Jennings, J. A. (2002). *An investigation of the validity of AP® grades of 3 and a comparison of AP and non-AP student groups* (College Board research report no. 2002-9). New York, NY: The College Board.
- Hargrove, L., Godin, D., & Dodd, B. G. (2008). *College outcomes comparisons by AP® and non-AP high school experiences* (College Board research report no. 2008-3). New York, NY: The College Board.
- Keng, L. & Dodd, B. G. (2008). *A comparison of college performances of AP® and non-AP student groups in 10 subject areas* (College Board research report no. 2008-7). New York, NY: The College Board.
- Kobrin, J. L., Patterson, B. F., Shaw, E. J., Mattern, K. D., & Barbuti, S. M. (2008). *Validity of the SAT® for predicting first-year college grade point average* (College Board research report no. 2008-5). New York, NY: The College Board.
- Market Data Retrieval, Inc. (2006). Spring 2006 school database. Shelton, CT: Market Data Retrieval.
- Mattern, K. D., Patterson, B. F., Shaw, E. J., Kobrin, J. L., & Barbuti, S. M. (2008). *Differential validity and prediction of the SAT®* (College Board research report no. 2008-4). New York, NY: The College Board.
- Morgan, R., & Crone, C. (1993). *Advanced Placement examinees at the University of California: An investigation of the freshman year courses and grades of examinees in biology, calculus AB and chemistry* (Unpublished ETS statistical report 93-210). Princeton, NJ: Educational Testing Service.
- Morgan, R., & Klaric, J. (2007). *AP® students in college: An analysis of five-year academic careers* (College Board Research report no. 2007-4). New York, NY: The College Board.

- Morgan, R., & Maneckshana, B. (2000). *AP® students in college: An investigation of their course-taking patterns and college majors* (Unpublished ETS statistical report 2000-09). Princeton, NJ: Educational Testing Service.
- Morgan, R., & Ramist, L. (1998). *Advanced Placement students in college: An investigation of course grades at 21 colleges* (Unpublished ETS statistical report No. 98-13). Princeton, NJ: Educational Testing Service.
- New Jersey Department of Education. (1999). *New Jersey world languages curriculum framework: A document in support of the core curriculum content standards for world languages* (pp. 45–46). Trenton, NJ: New Jersey Department of Education.
- Raudenbush, S. W., Bryk, A. S. (2002) *Hierarchical linear models applications and data analysis methods (advanced quantitative techniques in the social sciences series, second edition)*. Thousand Oaks, CA: Sage Publications, Inc.
- Shaw, E. J., & Mattern, K. D. (2010). *Examining the accuracy of self-reported high school grade point average*. Paper presented at the annual meeting of the American Educational Research Association, Denver, CO.
- Willingham, W.W., & Morris, M. (1986). *Four years later: A longitudinal study of Advanced Placement students in college* (College Board Research report no. 86-2). New York, NY: The College Board.

Footnotes

1. Dual enrollment refers to students completing college courses while still in high school.
2. Advanced Placement Italian Language and Culture was introduced in the 2005-06 academic year.
3. Note that the covariance parameter estimates in Table 5 may be compared directly with those reported in subsequent tables. Such comparisons demonstrate the extent to which the inclusion of student-level predictors led to the reduction in variation in SGPA that was attributed to either high school or college attended in the null model.

Table 2**Student Characteristics by Subject Area Sample**

Student Characteristic	Math	Computer Science	Engineering	Natural Science	Social Science	History	English	World Language	Arts & Music
Total Students	101,120	18,253	13,214	91,596	115,324	48,729	102,375	39,618	52,677
Gender									
Male ^a	48.9%	63.4%	80.8%	47.3%	44.0%	47.6%	45.1%	36.9%	43.3%
Female	51.1%	36.6%	19.2%	52.7%	56.0%	52.4%	54.9%	63.1%	56.7%
Racial or Ethnic Identity									
American Indian	0.4%	0.4%	0.3%	0.5%	0.5%	0.5%	0.4%	0.4%	0.4%
Asian / Pacific Islander	10.0%	9.9%	12.4%	10.7%	9.2%	7.0%	8.8%	9.4%	7.9%
Black / African American	7.2%	6.7%	4.1%	5.8%	7.1%	6.6%	7.8%	7.1%	6.3%
Hispanic / Latino	7.1%	5.2%	6.1%	6.8%	7.3%	8.1%	6.8%	7.1%	7.0%
Other	2.9%	2.8%	2.3%	2.9%	3.0%	2.9%	3.1%	3.6%	3.0%
White ^a	72.1%	74.7%	74.5%	73.0%	72.6%	74.6%	72.7%	72.1%	75.1%
Missing	0.3%	0.3%	0.2%	0.2%	0.3%	0.3%	0.3%	0.3%	0.3%
Best Language									
Another Language	1.0%	1.1%	1.2%	1.0%	0.9%	0.7%	0.9%	0.8%	0.8%
English & Another	4.7%	4.0%	4.4%	4.6%	4.4%	4.0%	4.4%	4.9%	3.8%
English Only ^a	91.3%	91.8%	91.7%	91.6%	91.6%	92.4%	91.5%	90.9%	92.3%
Missing	3.0%	3.0%	2.7%	2.8%	3.1%	2.9%	3.2%	3.4%	3.1%
Highest Parental Education Level									
H.S. Diploma or Less	21.6%	21.4%	15.6%	20.4%	21.5%	23.1%	22.4%	17.4%	20.3%
Associate Degree	6.4%	6.6%	5.4%	6.1%	6.3%	6.6%	6.5%	5.1%	6.2%
Bachelor Degree ^a	32.7%	33.3%	35.5%	33.2%	32.5%	33.4%	32.2%	31.0%	33.3%
Graduate Degree	32.2%	31.4%	36.9%	33.6%	32.6%	30.3%	31.7%	38.8%	33.3%
Missing	7.1%	7.4%	6.6%	6.7%	7.1%	6.6%	7.3%	7.7%	6.8%

^a Reference group for the relevant variable when used in subsequent analyses.

Table 3

High School Characteristics Weighted by Students in Each Subject Area Sample

High School Characteristic	Math	Computer Science	Engineering	Natural Science	Social Science	History	English	World Language	Arts & Music
Total Students	101,120	18,253	13,214	91,596	115,324	48,729	102,375	39,618	52,677
Affiliation									
Catholic	10.2%	11.4%	9.9%	9.5%	9.9%	10.1%	10.6%	11.9%	8.9%
Private, Non-Catholic	6.0%	4.6%	5.0%	6.4%	6.9%	5.9%	6.8%	9.6%	6.4%
Public	83.7%	83.9%	85.0%	84.1%	83.3%	84.0%	82.6%	78.5%	84.6%
Urbanicity									
Urban	27.6%	24.1%	27.9%	27.8%	28.0%	28.4%	27.3%	29.4%	27.8%
Suburban	53.9%	57.4%	55.2%	53.4%	53.9%	52.9%	54.6%	54.2%	54.5%
Rural/Non-Metropolitan	17.4%	17.2%	15.8%	17.7%	17.0%	17.6%	17.0%	15.3%	16.8%
Missing	1.1%	1.2%	1.1%	1.1%	1.1%	1.1%	1.1%	1.1%	1.0%
AP Exams Offered ^a									
0	2.1%	1.6%	1.0%	2.0%	2.1%	2.2%	2.1%	1.4%	1.9%
1-5	13.4%	12.7%	10.1%	13.3%	13.2%	13.4%	13.8%	11.3%	12.8%
6-10	23.4%	24.3%	20.9%	22.4%	23.2%	22.3%	24.3%	22.3%	23.1%
11-15	23.3%	23.7%	22.8%	22.9%	23.5%	22.8%	23.6%	24.0%	23.7%
16-20	19.5%	20.5%	21.6%	19.8%	19.8%	19.6%	19.5%	21.5%	19.8%
21 +	18.3%	17.1%	23.8%	19.5%	18.2%	19.6%	16.7%	19.5%	18.6%
Average Enrollment per Grade									
1-125	12.1%	9.2%	8.2%	12.1%	12.7%	12.1%	12.6%	14.0%	11.8%
126-225	14.7%	15.9%	13.0%	14.1%	14.5%	14.2%	15.2%	14.6%	14.3%
226-350	20.5%	22.4%	19.4%	19.9%	20.7%	19.4%	21.4%	20.9%	20.7%
351-499	24.2%	26.0%	25.6%	24.5%	24.4%	24.2%	24.9%	24.5%	24.9%
500 +	28.5%	26.5%	33.8%	29.4%	27.7%	30.1%	25.9%	26.0%	28.4%

^a Number of AP Exams with at least 5 examinees in a single year while the cohort attended high school (2003-06).

Table 4**College Characteristics Weighted by Students in Each Subject Area Sample**

College Characteristic	Math	Computer Science	Engineering	Natural Science	Social Science	History	English	World Language	Arts & Music
Total Students	101,120	18,253	13,214	91,596	115,324	48,729	102,375	39,618	52,677
Control									
Private	27.1%	27.0%	25.6%	25.1%	29.6%	25.3%	31.3%	44.0%	27.3%
Public	72.9%	73.0%	74.4%	74.9%	70.4%	74.7%	68.7%	56.0%	72.7%
Admittance Rate									
under 50%	17.4%	12.8%	19.3%	17.7%	18.0%	14.0%	18.1%	27.5%	17.7%
50 to 75%	56.6%	57.6%	50.9%	54.6%	56.4%	57.2%	56.1%	50.8%	56.8%
over 75%	26.0%	29.5%	29.8%	27.7%	25.6%	28.8%	25.8%	21.6%	25.5%
Undergraduate Enrollment									
750–1,999	3.3%	3.2%	0.9%	3.4%	4.0%	3.8%	4.4%	6.5%	4.1%
2,000–7,499	19.3%	23.1%	14.2%	15.3%	20.2%	20.4%	21.4%	27.7%	19.8%
7,500–14,999	28.0%	37.2%	24.1%	26.9%	27.4%	25.3%	29.8%	25.3%	29.6%
15,000 +	49.4%	36.5%	60.8%	54.4%	48.3%	50.5%	44.4%	40.4%	46.6%

