10-23-2008

The Relationship between AP English Language Performance and College Outcomes

Xinhui Xiong
Fordham University, Xiong_xinhui@yahoo.com

K. D. Mattern
College Board, kmattern@collegeboard.org

E. J. Shaw
College Board, eshaw@collegeboard.org

The Relationship between AP English Language Performance and College Outcomes

The Advanced Placement (AP) program, administered by the College Board since 1955, offers rigorous, college-level curricula and assessments at high schools across the United States and the world. It is viewed as a "cooperative educational endeavor" among high schools, colleges and universities. The AP program currently has standards for over 30 courses, including Art, History, Biology, Calculus, and English. The program offers students a unique opportunity to take more advanced courses during high school. Furthermore, each course has an end of year examination. Students who perform well on the examination may receive college credit or course exemption depending on the AP policies of the college or university they attend. Finally, there is a general belief that participation in AP courses helps students better prepare for the more demanding workload in college. As such, there has been a great deal of research devoted to examining AP performance and subsequent college outcomes (Ewing, 2006).

Research on AP Performance and College Success

Given the purpose of the AP program, it is not surprising that the majority of AP validity research has focused on the relationship between AP performance and course placement (e.g., Burnham and Hewitt, 1971; Dodd, Fitzpatrick, De Ayala, and Jennings, 2002; Klopfenstein and Thomas, 2006; Morgan and Crone, 1993; Morgan and Ramist, 1998). The results have generally found support for the AP program. Namely, students who perform well on an AP examination (a score of 3, 4, or 5) and receive course credit for the examination tend to outperform non-exempt students in subsequent courses, even after controlling for academic preparedness (e.g., standardized test scores).

Additional AP research has examined more general outcomes. For example, a study by Willingham and Morris (1986) found that students who took an AP examination, regardless of

performance, were more likely to earn a B average during their first year of college as compared to students who did not take any AP examinations. This was true even after controlling for academic ability. Furthermore, a study by Dougherty, Mellor, and Jian (2006) found that students who performed well (a score of 3 or higher) on at least one AP examination in English, mathematics, science, or social studies were more likely to graduate from college in five years as compared to students who either took no AP examinations, who received a score of 1 or 2, and also those who took an AP course but not the examination. Again, this was true even after controlling for academic ability and other student/school characteristics.

The purpose of this study is to build on the extant body of research highlighting the efficacy of the AP program. Specifically, the current study will examine the relationship between AP English Language performance and subsequent college success, as indexed by first-year college GPA (FYGPA), retention to the second year of college, and the selectivity level of the institution attended, after controlling for SAT performance. The AP English Language examination was selected for two primary reasons. First, it is one of the highest volume AP examinations thus ensuring sufficient data for analyses. Second, and more importantly, its content is relevant to college performance regardless of academic major. Analyzing data from 110 institutions for roughly100,000 students, this study represents the largest sample in AP validity research to date, thereby increasing the generalizability of the results as well as minimizing sampling error.

Method

Sample

The data analyzed in the current study are from the SAT Validity Study database (see Kobrin et al., 2008, for more information). This database is comprised of student level data for

first-time, first-year students in the entering class of 2006 at 110 participating colleges and universities in the U.S. Course-level performance data, FYGPA, and retention data were matched back to College Board databases to include SAT scores, SAT Questionnaire responses, AP scores, and institutional characteristics.

Students were then classified into three groups according to their AP English Language examination performance. Specifically, students who did not take any AP examinations were classified as Group 1. Students who took the AP English examination and received a score of 1 or 2 were classified as Group 2, and students who received a score of 3 or higher were classified as Group 3. Therefore, students who did not take AP English Language but took another AP examination were excluded from the current study. Furthermore, students without SAT scores were excluded from the analysis. This resulted in a final sample size of 92,964 students who had complete data on all study variables. Table 1 provides the distribution of these three groups.

Measures

AP English Language scores. Official AP English Language scores were obtained from College Board records. AP scores are criterion referenced and range from 1 to 5. A score of 1 represents '*No recommendation'*; 2 represents '*Possibly qualified'*; 3 represents '*Qualified'*; 4 represents '*Well-qualified'*; and 5 represents '*Extremely well-qualified'*.

SAT scores. Official SAT scores were obtained from College Board records. The SAT is composed of three sections: Critical Reading, Math, and Writing. The score scale for each section ranges from 200 to 800. The composite SAT score is the sum of the three sections scores, and ranges from 600 to 2400.

SAT Questionnaire. Gender, race/ethnicity, and best language spoken were self-reported by students on the SAT Questionnaire, which is completed at the time of SAT registration.

Race/ethnicity was collapsed into seven categories: American Indian or Alaska Native; Asian, Asian-American, or Pacific Islander; Black or African-American; Hispanic; White; Other; and No Response. Best language spoken was classified into four categories: English only, English and another language, Another language, and No Response.

First-Year College GPA (FYGPA). FYGPA was supplied by participating institutions and ranged from 0.00 to 4.27.

Retention to the second year. Participating institutions indicated whether students who entered in the fall of 2006, returned for the second year of college in fall 2007. Students who did return for the second year received a value of 1, whereas students who did not return received a value of 0.

Institution Selectivity. Institution selectivity is the percentage of applicants that were admitted to the institution. The higher the percentage of students admitted by an institution, the less selective it is considered to be. These percentages were computed from institution responses to the College Board's Annual Survey of Colleges.

Analyses and Results

Descriptive Statistics

The demographic characteristics of the three AP groups are provided in Table 2. Female students outnumbered male students within each group: 53% versus 47% for Group 1, 62% versus 38% for Group 2, and 58% versus 42% for Group 3. As for race/ethnicity, White students comprised the majority within each group: 69% in Group 1, 54% in Group 2, and 67% in Group 3. However, minority students, namely Hispanic and African-American students, made-up a significantly larger proportion of Group 2 as compared to Group 3. Students who stated that English was their best language represented the majority of each group: 91% in Group 1, 88% in

Group 2, and 93% in Group 3. However, similar to the race/ethnicity results, students reporting that their best language was not English represented a larger proportion of Group 2 as compared to Group 3.

Table 3 provides the mean FYGPA, retention rate to the second year, institution selectivity rate, and SAT score for each group. The results indicate that Group 1 had the lowest mean SAT score (1539) whereas Group 3 had the highest mean SAT score (1933). Furthermore, Group 1 had the lowest mean FYGPA (2.74), lowest second year retention rate (80%), and attended the least selective institutions (68% of applicants were admitted). Group 3 had the highest mean FYGPA (3.30), highest second year retention rate (92%), and attended the most selective institutions among the three groups (56% of applicants were accepted).

Predictive Validity

For the two dependent variables (outcomes) of FYGPA and institutional selectivity of college attended, after checking the linear trend, homogeneity of the variance, homoscedasticity, data were analyzed using ANCOVAs with AP English Language performance group as the independent variable (predictor) and SAT composite score entered as a covariate to control for academic ability. There was a small interaction between group 1 and group 2, but not serious. Additionally, retention to the second year was predicted from AP English Language group, controlling for SAT composite, with logistic regression after checking the homoscedasticity. Contrasts were computed for all possible group comparisons.

Table 4 provides the results of the group contrasts for FYGPA, institutional selectivity, and retention to the second year without controlling for SAT composite score. All group differences are statistically significant. Specifically, students who took AP English Language and scored a 3, 4, or 5 performed significantly better on all three academic outcomes as compared to

students who scored a 1 or 2 and to students who didn't take any AP examinations. Moreover, students who took AP English Language and scored a 1 or 2 performed significantly better on all three academic outcomes as compared to students who didn't take any AP examinations. However, because students are not randomly assigned to AP classes, the student's academic ability should be taken into account in order to disentangle the effects of AP performance on future academic outcomes from academic achievement (e.g., SAT scores).

Table 5 provides the results of the group contrasts for FYGPA, institutional selectivity, and retention to the second year controlling for SAT composite score. The differences in academic outcomes across groups are smaller when controlling SAT composite scores; however, they remain statistically significant, except for the difference in institutional selectivity between Group 1 and Group 2. Specifically, the mean FYGPA of Group 1 was 0.06 lower than that of Group 2, the mean FYGPA of Group 2 was 0.12 lower than that of Group 3, and the mean FYGPA of Group 1 was 0.18 lower than that of Group 3. The mean institutional selectivity (percentage of applicants admitted) of Group 1 was 0.2% (not significant) higher than that of Group 2, the mean institutional selectivity of Group 2 was 4.6% higher than that of Group 3, and the mean institutional selectivity of Group 1 was 4.8% higher than that of Group 3. That is, students in Group 1 and Group 2 were accepted by institutions of approximately the same selectivity level.

As for retention to second year, the difference in retention rates was significantly different for each pair of groups. From the odds ratio estimates (ratios of odds of lower ranked group to that of higher ranked group), students in Group 1 had the lowest chance of returning to the school for a second year, and students in Group 3 had the highest chance of returning. After controlling for students' SAT composite scores, the same trend maintained, but the odds ratios

all increased. The increased odds ratios means that controlling for SAT reduced the difference in retention rates among groups but did not eliminate it since all paired contrasted remained significant.

Discussion and conclusion

This study demonstrated that higher AP English Language examination performance corresponded to higher FYGPA and second-year retention rates. Students with the highest AP English Language scores (scores of 3 or higher) also attended slightly more selective institutions. Even after controlling for SAT composite score, the same pattern of results remained with the exception of the selectivity of institution attended by the student.

While association is not the same thing as causation, the results of this study provide some support for the role of the AP program in subsequent college performance and success. Though students with stronger academic backgrounds are more likely to participate in the AP program, earn higher AP scores and FYGPAs, have higher second-year retention rates, and attend more selective institution, this study showed that even when prior academic performance was controlled for, significant group differences still existed for the AP English Language performance group comparisons. That is, after controlling for the effects of prior academic performance, those with a 3, 4, or 5 on the AP English Language examination tended to outperform students who received a 1 or 2 on the AP exam, as well as students who did not take the AP exam, with regard to FYGPA. These results suggest that participation in the AP program may better prepare students for the more rigorous study schedule in college. Nevertheless, it is possible that other factors beyond prior academic performance contribute to the group differences. Future research should identity other useful variables to control for when examining the impact of AP performance on academic outcomes. Additionally, this study only

examined the effects of one AP examination. It would be useful to determine whether the same pattern of results hold for other AP examinations.

Future research

Based on the caveats described above, there are several avenues for future research.

First, the analyses should be replicated with other AP examinations, particularly outside the English content area. This would test whether this pattern of group differences generalizes to all AP examinations, or is unique to English Language examination. For example, the difference between Groups 1 and 2 on institutional selectivity was not statistically significant. Future research should test whether this pattern holds across examinations. If it does, it would be interesting to explore why and how a student's performance on an AP examination may influence the type(s) of colleges to which he/she applies, is admitted, and ultimately enrolls.

Secondly, additional outcomes should be examined such as college-going rates, cumulative GPA, and graduation rates. For example, it would useful to understand whether there are differences in the percentage of students attending college among the three AP performance groups, as well as by AP examination area. Cumulative GPA could be assessed as an outcome to determine whether the initial benefit of AP performance carries through to more distal college outcomes. Similarly, graduation is the ultimate goal of college and should be regarded as one of the more important measures of college success. Therefore, future research should also determine the relationship between AP participation and graduation.

Thirdly, other student characteristics, such as parental education and income, HSGPA, and high school characteristics should be examined when analyzing the relationship between AP performance and subsequent college success. That is, does the AP effect remain once these other variables are also considered?

Finally, school characteristics, such as public or private, could also be taken into consideration. Using hierarchical linear modeling, difference between schools versus within schools could be explored more to assess the effect of school factors on the relationship between AP performance and subsequent college success.

References

Burnham, P. S., & Hewitt, B. A. (1971). Advanced Placement scores: Their predictive validity. *Educational and Psychological Measurement, 31*, 939–45.

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: The College Board.

Dougherty, C., Mellor, L., & Jian, S. (2006). The relationship between Advanced Placement and college graduation. (National Center for Educational Accountability: 2005 AP Study Series, Report 1). Austin, Texas: National Center for Educational Accountability.

Ewing, M., (2006). *The AP Program and Student Outcomes: A Summary of Research.* (College Board Research Note RN-29). New York: The College Board.

Kobrin, J. L., Patterson, B. F., Shaw, E. J., Mattern, K. D., & Barbuti, S. M. (2008). *Validity of the SAT Reasoning Test for Predicting First-Year College Grade Point Average*. (College Board Research Report No. 2008-05). New York: The College Board.

Klopfenstein, K., & Thomas, M. K. (2006). The link between Advanced Placement experience and early college success. Retrieved July 1, 2006, from http://www.utdallas.edu/research/tsp/pdfpapers/ap_coll.060706.pdf.

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.* (ETS Statistical Report 93-210). Princeton, NJ: Educational Testing Service.

Morgan, R., & Ramist, L. (1998). *Advanced Placement students in college: An investigation of course grades at 21 colleges.* (ETS Statistical Report No. 98-13). Princeton, NJ: Educational Testing Service.

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: The College Board.

Table 1

Frequency and Percentages of the Three AP Performance Groups

AP Group	**N**	**Percent**
Group 1: Took no AP Examinations	60,955	65.6
Group 2: Took AP English Language and scored a 1 or 2	10,375	11.2
Group 3: Took AP English Language and scored a 3,4 or 5	21,634	23.3
Total	92,964	100.0

Table 2

Demographic Characteristics of the AP Performance Groups

Variable		Group 1: No AP	Group 2: AP Eng. (1,2)	Group 3: AP Eng. (3,4,5)
Gender	Male (%)	28,472 (46.7)	3,926 (37.8)	8,895 (41.1)
	Female (%)	32,483 (53.3)	6,449 (62.2)	12,739 (58.9)
Race/ Ethnicity	American Indian / Alaska Native (%)	371 (0.6)	59 (0.6)	95 (0.4)
	Asian/Asian-American/Pacific Islander (%)	3,869 (6.3)	1,202 (11.6)	2,482 (11.5)
	Black (%)	4,968 (8.2)	959 (9.2)	640 (3.0)
	Hispanic (%)	3,693 (6.1)	1,658 (16)	1,311 (6.1)
	White (%)	42,404 (69.6)	5,697 (54.9)	14,749 (68.2)
	Other (%)	1,696 (2.8)	286 (2.8)	634 (2.9)
	No Response (%)	3,954 (6.5)	514 (5.0)	1,723 (8.0)
Best Language	English Only (%)	55,487 (91)	9,166 (88.3)	20,096 (92.9)
	English and Another Language (%)	2,362 (3.9)	845 (8.1)	809 (3.7)
	Another Language (%)	772 (1.3)	102 (1.0)	40 (0.2)
	No Response (%)	2,334 (3.8)	262 (2.5)	689 (3.2)

Table 3

Mean Performance of Study Variables by AP Performance Groups

Variable	**Group 1: No AP**	**Group 2: AP Eng. (1,2)**	**Group 3: AP Eng. (3,4,5)**
SAT	1539	1618	1933
FYGPA	2.73	2.87	3.3
Retention	0.83	0.88	0.93
Institution Selectivity	0.68	0.67	0.56

Note. Institution selectivity is the ratio of number admitted students divided by the number of applicants. Larger numbers indicate less selective institutions.

Table 4

Paired Contrasts for the AP Performance Groups

Variable	Contrast	Point Estimate	95% C.I. Lower	95% C.I. Upper	Sig.
FYGPA	No AP vs. AP Eng. (1,2)	-0.137	-0.151	-0.122	0.000
	AP Eng. (1,2) vs. AP Eng. (3,4,5)	-0.425	-0.441	-0.408	0.000
	No AP vs. AP Eng. (3,4,5)	-0.562	-0.573	-0.551	0.000
Institution Selectivity	No AP vs. AP Eng. (1,2)	0.012	0.008	0.015	0.000
	AP Eng. (1,2) vs. AP Eng. (3,4,5)	0.103	0.099	0.106	0.000
	No AP vs. AP Eng. (3,4,5)	0.114	0.112	0.117	0.000
Retention	No AP vs. AP Eng. (1,2)	0.645	0.606	0.687	0.000
	AP Eng. (1,2) vs. AP Eng. (3,4,5)	0.536	0.495	0.580	0.000
	No AP vs. AP Eng. (3,4,5)	0.346	0.327	0.366	0.000

Note. Point estimate for retention, is in odds ratio units. It is the ratios of odds of lower ranked group to that of higher ranked group

Table 5

Paired Contrasts for the AP Performance Groups with SAT as a Covariate

Variable	Contrast	Point Estimate	95% C.I. Lower	95% C.I. Upper	Sig.
FYGPA	No AP vs. AP Eng. (1,2)	-0.060	-0.074	-0.046	0.000
	AP Eng. (1,2) vs. AP Eng. (3,4,5)	-0.121	-0.138	-0.104	0.000
	No AP vs. AP Eng. (3,4,5)	-0.181	-0.194	-0.168	0.000
Institution Selectivity	No AP vs. AP Eng. (1,2)	-0.002	-0.005	0.001	0.224
	AP Eng. (1,2) vs. AP Eng. (3,4,5)	0.049	0.046	0.053	0.000
	No AP vs. AP Eng. (3,4,5)	0.047	0.045	0.050	0.000
Retention	No AP vs. AP Eng. (1,2)	0.720	0.676	0.767	0.000
	AP Eng. (1,2) vs. AP Eng. (3,4,5)	0.813	0.747	0.884	0.000
	No AP vs. AP Eng. (3,4,5)	0.585	0.547	0.626	0.000

Note. Point estimate for retention, is in odds ratio units. It is the ratios of odds of lower ranked group to that of higher ranked group