

An Argument for Changing Institutional Policy on Granting AP Credit in English: An Empirical Study of College Sophomores' Writing

Kristine Hansen, Suzanne Reeve, Richard Sudweeks, Gary L.
Hatch, Jennifer Gonzalez, Patricia Esplin, William S.
Bradshaw

INTRODUCTION

Granting credit, advanced placement, or both to incoming college students on the basis of performance on Advanced Placement (AP) exams has long been an accepted and common practice at many colleges and universities. In 2002 just under one million high school students (most of them college-bound juniors and seniors) took nearly 1.6 million AP exams (College Board, *AP Central*). Assuming most of those students received credit at the institutions where they eventually matriculated, it is clear that, by design or default, AP programs have become a de facto part of college curricula. In the case of English, by far the largest of the AP programs, this fact has important implications for writing program administrators, who have no influence over courses that are offered elsewhere, taught by staff they did not hire or train, and evaluated by means of tests they had no hand in designing. Every WPA must ask, along with Mahala and Vivion, “Do the [AP English] exams that lead to the granting of credit reflect mastery of the same knowledge, the same critical thinking abilities, and the same academic competencies as the courses for which they substitute?” (43).

These questions have concerned us at Brigham Young University (BYU), where since the 1962–63 academic year, incoming students have been awarded between six and twelve credit hours (including three for first-year writing) for scores of 3, 4, or 5 on either or both of the AP English exams—Literature and Composition or Language and Composition. BYU is a selective, private, religiously affiliated university enrolling approximately thirty-two thousand students from all fifty states and 130 foreign nations. The average incoming student has a high-school GPA of 3.76 and an ACT composite score of 27. BYU is ranked among the top ten universities for numbers of students entering with AP credit (CB, *The 200 Colleges and Universities*). Thirty-five to 40 percent of each first-year class in the past five years has entered with AP English scores of 3, 4, or 5. Half of those students have earned the score of 3.

First-year students at BYU have the option of enrolling in one of four first-year composition (hereafter called FYC) courses: English 115, College Writing and Reading; English 200, Rhetoric and Writing; Honors Program 200, Honors University Writing; or Philosophy 105, Reasoning and Writing. English 115 is regarded as the standard course, enrolling about 60 percent of students who take FYC. The other courses are considered advanced first-year courses, designed for students who plan to seek honors graduation or who want a challenge different from the standard course. There is no placement exam; students are simply advised to choose the course they think best meets their needs. About 40 percent of the entering students with AP English scores of 4 or 5 choose to enroll in FYC, usually Honors 200. In contrast, only about 30 percent of students with AP scores of 3 choose to enroll in FYC.

Whether students take or bypass FYC, they are still required to take one of thirteen advanced writing courses, most of which focus on writing in the disciplines. Although students are advised to take advanced writing early in their junior year, many put off doing so until the last semester of their senior year. As a result, 30 percent to 40 percent of the students in every entering class take no university writing course until they are juniors or, more likely, seniors.

Given these facts, we have questioned the wisdom of BYU's forty-year-old policy on granting AP English credit.¹ We have been particularly concerned with the level of achievement of students who earn a score of 3, between eight hundred and one thousand students in every entering class. According to the grade scale published by the College Board, a 3 indicates that a student is "qualified" for advanced placement—which, at BYU, means he or she receives credit for FYC. Furthermore, the AP guide explains, "Generally, to obtain a grade of 3 or higher on an AP Exam, students need to

answer about 50 percent of the multiple-choice questions correctly and do *acceptable* work on the free-response section” (College Board, *Access to Excellence*, emphasis added). The implications of defining 50 percent on a multiple-choice exam as “qualified” for advanced college work are troubling, as is James Vopat’s statement that “acceptable” work on the essays is writing “formally defined as lacking in detail, support, and appropriate focus” (58).

Writing program administrators and teachers at BYU (many of whom have participated in scoring AP English exams) agree that students earning a 3 would benefit substantially from taking FYC. Since at least 1990, WPAs and chairs in the English department have argued strenuously that credit should no longer be awarded for scores of 3. However, for several reasons, not the least of which is the cost of adding about forty sections of FYC, the university administration has been reluctant to change the policy without convincing data indicating that a change is necessary.²

In response to efforts to raise the AP bar, university administrators have cited the results of two grade studies completed in 1994 and 1998 by BYU and College Board personnel (Morgan and Ramist). These studies concluded there was no difference in the average grades of advanced writing students who had taken FYC and those who had bypassed it with AP credit. An internal study completed in 2000 corroborated the earlier findings. As a result, university administrators argued that FYC apparently added no value to students’ educations and its absence did no harm to AP students; therefore, they found no empirically valid reason to raise the standard for awarding credit for AP tests.

Beyond the comparative grade studies conducted at BYU, there is little substantial research——particularly independent research——focusing on the predictive validity of AP English scores for student success in college. This is an important issue in light of the prevalence of AP English courses and recent expansion of AP programs. Few studies have been performed during the last ten years, and some are twenty years old or more.

The studies that exist fall into three main categories. The first examines correlations between AP participation and performance and overall college performance as measured by time to graduation, courses taken, academic achievement (for example, GPA), and the like. The findings of these studies vary. For example, Chamberlain et al., Creech, and Willingham found that AP students complete more credit hours per semester, have a higher percentage of upper-division credits, and have a higher overall GPA than non-AP students. Additionally, Creech concluded that students who take AP courses (regardless of whether they receive college credit) perform better in college

than those who do not participate in AP. Cusker and LeMy's work indicated that students earning AP credit were more likely to pursue more ambitious programs (such as double majors and additional minors).

The second type of studies compares performance of high school and college students on AP exams and is the only method that includes a measure of writing; essays are part of the exam, and they are assessed using only the exam rubric. Of the three studies in this category of which we are aware, two compared various samples of high school students (Bodenhausen; Dvorak), while one compared AP high school students with college students in equivalent courses (Modu). We have found no published studies evaluating predictive validity based on assessment of essay writing not associated with the AP exam.

The third category, exemplified by Morgan and Ramist's research, evaluated validity of individual exams by comparing grades of AP and non-AP students in subsequent English courses. Richardson's 1978 study looked specifically at AP English scores and grades in the first college English course taken. More recently, Bridgman and Lewis compared grades in a variety of subjects, including English. All of these studies concluded that AP English exam scores correlated significantly with grades in equivalent college English courses, and that AP English participants received higher grades than non-participants in subsequent English courses (see also Mercurio; Mercurio et al.; Burnham and Hewitt). This method is attractive for its relative simplicity and is recommended in the AP literature as a means of testing validity at individual schools. However, our research team found that the reliance on the results of grade studies by our university's administration suspect for four reasons.

First, FYC and advanced writing courses at BYU are not designed to be contiguous or even to build directly on each other, so the grade in a second course should not be considered a measure of learning that did or did not take place in earlier writing courses. Second, 80percent to 90 percent of students in advanced writing at BYU score an A or a B, making the performance range so small that differences are almost impossible to detect. (In fact, the 2000 grade study found no difference among AP students who scored 3, 4, or 5.) Third, criteria for assigning grades vary from teacher to teacher and include, in addition to performance on writing tasks, variables such as attendance, class participation, reading quizzes, and completion of homework exercises. Thus, grades are a highly suspect means of comparison; what would be an A in one teacher's class might be a B in another's. Fourth, because most students take advanced writing three to six years after entering BYU, it is likely that any earlier differences that were present between AP and non-AP students have "flattened out," either because both groups have

forgotten some of what they knew about writing or because intermediate learning, maturation, and experience have helped students compensate for writing deficiencies that might have been present earlier.

These arguments about the invalidity of grade studies have not convinced BYU's administration to change the English AP policy. Additional arguments from William Lichten's impressive study of AP growth and exam scoring, however, support the inadvisability of giving credit for the AP score of 3. Lichten calculated that about 45 percent of four-year institutions—particularly among selective and highly selective institutions—no longer award credit for AP scores of 3 because the level of achievement signified by the score of 3 has declined. He attributes the decline to several interrelated factors. First, the pool of AP test-takers has grown to five hundred times its original number in 1953, a fact that inevitably means less-qualified students are now involved in AP programs and taking the tests. Second, AP exams are quota-based. That is, final scores are calculated so that two-thirds of those taking the test will score 3, 4, or 5. This method of calculation is justified when a consistently qualified group of students takes the test each year, because it ensures that variation in the difficulty of exams from year to year does not compromise the consistency of evaluation and student performance. However, when the pool expands to include many who are only marginally qualified, it inevitably means that more students of lesser ability are included in the quota. The third reason for declining quality, according to Lichten, is that state legislatures have mandated higher participation in AP programs on the assumption that the states will save money if students earn college credit while still in high school. Some states even pay the cost of the tests and urge students who have not even taken the AP course to take the test anyway. All of these factors have led, in Lichten's appraisal, to a situation fraught with irony: At a time when so-called college courses are proliferating in high schools, colleges find themselves offering more remedial instruction than ever.

Some of Lichten's findings have been disputed by the College Board's Wayne Camara in a response co-authored by Neil J. Dorans, Rick Morgan, and Carol Myford (three employees of Educational Testing Services, the agency that conducts the scoring of the AP exams). Significantly, however, they do not directly address the charge that less qualified students may now be earning a score of 3. Lee Jones, executive director of the AP Program, acknowledged in a response to Lichten's study that the increased number of students taking AP exams has resulted in the inclusion of students who may be less talented or motivated, but he describes the expansion as an effort to increase opportunities for traditionally excluded students to take AP courses (cited in Ganeshanathan A45).

We cannot judge whether the apparent trend in universities of not accepting scores of 3 for credit is due to declining quality in performance on the exams or to changes in college requirements; nevertheless, it seems likely that the score of 3 does not represent a desirable level of achievement, which in turn suggests that an AP policy that is now forty years old should be carefully examined.

Responding to these arguments, an associate academic vice-president at BYU stated he would welcome additional data that would help determine whether the bar for granting credit and exemption from FYC should be raised. As we designed the study reported here, we determined that we should study the writing of students early in their college career. We believed students with AP scores of 3 would feel most keenly the effects of having missed instruction in college writing during or soon after the freshman year, as they tried to complete writing assignments in other courses for which their high school skills were inadequate. Although differences between AP and non-AP students might not be evident in the senior year, that should not justify a policy that may contribute unnecessarily to students' floundering about in the first two or three years of college. Accordingly, we focused the present study on sophomore students, defined by credit hours in residence at the university. We also determined not to use grades as the dependent variable, but to use direct measures of students' writing ability. We compared the writing of three groups: (1) students who had AP credit but nevertheless took a first-year writing course; (2) students who had AP credit and chose to bypass first-year writing; and (3) students who did not have AP credit and therefore took a first-year writing course.³

Methods

PARTICIPANTS

Four teachers agreed to cooperate in the study, all of whom taught a sophomore-level general education course on the history of civilization. One of the teachers taught a very large section (approximately 180 students) of Humanities 201. Another taught two small sections of Humanities 201 (40 students each) and a large section (180 students) of History 201. Two of the teachers taught small honors sections of Humanities 201 (30 to 35 students each). These classes were chosen because we knew they were likely to have sufficiently large enrollments of sophomores in the three groups we were interested in. The honors sections were included because the BYU Honors Program does not allow exemptions from first-year writing regardless of AP credit, and we wanted to study the writing of some students who had taken first-year writing even though their AP scores would have otherwise allowed

them to bypass the requirement. Composition courses themselves were excluded because we wanted to study the *effects* of having taken or bypassed FYC; additionally, we wanted to analyze writing produced in an academic setting where formal writing instruction was not given.

All of the teachers agreed to allow us to write the prompts for two papers they would assign, and they agreed to give the students course credit for these papers so that students would be motivated to do their best. The amount of course credit to be awarded was left to each teacher's discretion. The students understood that these papers would be used in a research study; however, they were also informed that their teachers had helped design the assignments and would also read what they had written. (A comparison of mean essay scores between sections indicated that the methods and personalities of the teachers were not confounding variables.) In all, the history of civilization classes enrolled 497 students, among whom were 214 sophomores—students with between two and four semesters in residence at BYU. Of these sophomores, 182 (8 percent) wrote both essays and returned all required materials. These 182 students became the major focus of the study.

MATERIALS

Two writing prompts were created with the advice and approval of the teachers involved. The first prompt (see Appendix A) was based on Arthur Miller's "Tragedy and the Common Man." This prompt asked students to read Miller's essay and, with reference to a tragic text they had read in class, discuss whether they would agree with, disagree with, or modify Miller's claim that "tragedy implies more optimism in its author than does comedy." The second prompt (see Appendix B) was based on an essay entitled "The Need Beyond Reason," by Edward Hart. This prompt asked students to agree with, disagree with, or modify Hart's claim that the value of the humanities lies precisely in their non-utilitarian nature, and to use as evidence for their position examples from their study of the humanities in the course in which they were enrolled.

The research team also created a scoring rubric for each prompt, using a nine-point scale and descriptors of the kind of essay that would merit each score on the scale (see Appendices C and D). Readers of this article who have participated in the summer scoring of AP essays will recognize that the rubrics are very similar to those employed in the AP scoring. The writing prompts were also somewhat similar to those used in the AP exams. This imitation was deliberate on our part. While we knew this decision might actually favor the students who had taken AP courses, we were willing to run that risk because we believed that the prompts asked students to read and think critically and to produce a kind of writing that is commonly assigned

at the university. More importantly, however, a prompt of this design served as a kind of control: it helped eliminate differences between prompts used in the AP exam and those used in our study as a confounding variable. Differences in writing performance could then more easily be interpreted as being due to the poor predictability of the AP exam, not to differences in the essay assignments.

In addition to the prompts and rubrics, the research team devised or adapted several instruments to collect additional information from the students: (1) a writing process questionnaire that students filled out upon submission of their essays to indicate what they did in planning and producing each essay; (2) an adapted version of the writing self-efficacy questionnaire developed by Shell et al. to measure students' confidence in their ability to write in various genres; (3) the Daly-Miller Writing Apprehension survey; and (4) a general questionnaire about students' past experiences writing at home and in school. (We plan to report the results of these measures in a subsequent article.)

PROCEDURES

For each prompt, students were given one week to write a three-page, word-processed essay and submit it to their teacher; the preparation period was the same for every participant regardless of course or section. The virtue of this format is that we probably gained an authentic sample of each student's writing ability. The potential weakness is that some students may not have done their own work. Our best evidence that this was not a significant problem comes from the students' responses to our questionnaire (mentioned above) about the process they used in writing: it was rare that students reported even having asked another person to read a draft of their work. Our strong impression is that the essays were written without collaboration.

Teachers assigned the first paper in late September or early October and the second paper in late October or early November of 2002. The research team collected and photocopied the essays for later scoring. Identification numbers were given to each essay so that students would remain anonymous. Two members of the research team read dozens of essays to assess the range of performance on each prompt and to find suitable essays to use in training raters.

Twelve raters were chosen from the ranks of experienced adjunct faculty and graduate students who teach first-year and advanced writing. They were trained on two separate occasions using essays representing a range of performances on each prompt. A generalizability analysis was conducted to assess the reliability of the ratings and to determine to what extent the reliability was decreased by (1) interrater inconsistencies, (2) intrarater inconsistencies,

(3) intratask inconsistencies (differences in the relative difficulty of the writing prompts), or (4) the various interactions among these potential sources of extraneous variance (see Brennan; see also Cronbach, Gleser, Nanda, and Rajaratnam). An article by Sudweeks, Bradshaw, and Reeve describing the results of the generalizability analysis is in preparation.

After obtaining satisfactory measures of interrater reliability (mean intraclass $r = .90$ following Shrout and Fleiss, equation ICC (2,k) p. 426), the scoring began. All essays written by nonsophomores were rated once. The 364 essays written by the 182 students in the sophomore group were each rated once by two different readers. For forty-two essays (about 11 percent) the judgments of the first two readers varied by three or more points (for example, 5 and 8). These essays were read an additional time (total of three independent readings), and outlying scores were discarded. For purposes of this study, the independent readers' scores were averaged on each essay; these scores were then averaged again across essays to create one score for each student. No scores were rounded, so no variation in scores was introduced by the research team.

RESULTS

Distribution of freshmen in composition course offerings

Table 1 shows the percentages of first-time freshmen at BYU in 2000–01 who completed FYC through either the standard course (English 115) or honors/advanced FYC. As shown, the freshman class in this year consisted of approximately four thousand students, over one-third of whom entered with AP English credit. A majority of the AP students (60 percent) opted out of the composition requirement during their freshman year, in spite of university advisement urging them to enroll. The remaining AP students overwhelmingly chose to enroll in honors or advanced freshman composition offerings. Among students not awarded AP English credit at entrance to the university, nearly 70 percent enrolled in English 115, 20 percent enrolled in honors or advanced courses, and the remainder postponed meeting the requirement.

Essay scores by AP English status

Table 2 shows the results of a one-way analysis of variance comparing writing ability (mean essay score) among sophomore students grouped according to AP and composition course backgrounds. There appears to be a significant gain contributed by the experience in a university writing course; those students receiving AP English credit *and* completing a freshman composition course (AP + FYC) scored fully a point higher than either of the other two groups on the 9-point scale ($p < 0.005$). Scores from students with

AP English credit who bypassed the composition requirement (AP, no FYC), and students with no AP English credit who completed a freshman composition course (no AP, FYC) were not statistically different from each other.

Essay scores by AP English scores

Table 3 shows the results of a two-way analysis of variance between students who passed one or more AP English exams with a score of exactly three and students receiving a score of greater than three, subdivided by whether or not each student completed a freshman composition course. Because of the small numbers of students with AP scores of 4 and 5, these two groups were collapsed into one cell. In the two-by-two analysis, the differences between row means are statistically significant ($p = 0.002$) and reflect differences in mean writing ability between students who completed a freshman composition course and those who did not. The differences between the two column means are also significant ($p = 0.04$) and show that students who passed the AP English exam with a score greater than 3 write significantly better than those who obtained only a 3. The lack of a significant interaction between cells indicates that differences in the row means are generalizable across both AP conditions. The important insight derived from these data, then, is the tendency to low performance by students who scored 3 and bypassed the FYC requirement. Though the mean score of this group (4.91) is likely not significantly different from the group mean of 5.13 cited in Table 2 for students who lacked AP and took FYC, we suggest that both scores indicate an unacceptably inferior level of writing skill, whether achieved by students who have completed a freshman composition course or an AP course deemed to be its equivalent.

The students in this sophomore data set ($n = 182$) constitute about 5 percent of an entering class and are a representative sample of the population of interest in terms of GPA, ACT score, and percentage of students with AP English credit (about one-third; see Table 1). We acknowledge that the sample sizes seem small, but we remind readers of the constraints involved in obtaining a sample of even this magnitude. The relatively small sample sizes make the resulting significant p -values even more indicative of real differences between the groups surveyed.

Composition course enrollment by “best” and “poorest” writers

The data in Table 4 represent an attempt to confirm the validity of the foregoing conclusions by an alternative approach. Table 4 shows the distribution of composition courses taken by writers at the high and low skill extremes of our sample (high skill defined as students scoring 5 or above on both essay prompts, with at least one score of 8 or 9; low skill defined as students scoring 5 or below on both essay prompts, with at least one score

of 1, 2, or 3). A chi-square analysis confirmed that the course distributions of the two groups were statistically different ($p < 0.0004$). Clearly, a higher proportion of the best writers had enrolled in and completed an honors or advanced first-year composition course.

Essay scores by hours in residence

The scores obtained by 429 students in our sample (students for whom total credit hours in residence at a college or university could be determined) are displayed in Table 5. Credit hour listings are not official university statistics, but were obtained by subtracting from students' total credit hours any AP, high school concurrent enrollment, or other credits by examination not achieved through attendance on a college campus. The aim was to assess whether students' writing skill improved through the time spent and experience gained as a matriculated college student. As shown in the table, no significant difference in essay scores was observed between students according to class standing. These results may be biased by relatively small percentages of upperclassmen in the selected courses, as well as by the idiosyncratic scholastic characteristics of upperclassmen who procrastinate taking a required 200-level course until their senior year. More important, a direct measure of writing competence in students approaching graduation would require a research design different from that used here.

RESULTS

This study was undertaken after it became clear that, although the opinions of trusted faculty with regard to the granting of AP credit were appreciated, administrators would be reluctant to effect a change in the forty-year-old AP policy without strong empirical evidence that such a change was needed. They did not object to increasing the writing competence of our students, and they were willing to encourage students with AP credit to enroll in FYC. But they believed fiscal constraints did not permit adding some forty new sections of FYC without strong reasons to reallocate existing funds or request additional funds. This study was designed to determine whether we could provide the compelling evidence that would help administrators secure the funds.

A study of this kind is conducted under certain logistic constraints. Having made the decision to conduct the research within the framework of existing courses, we faced the primary challenge of selecting a set of classes that enrolled students who could all respond to the same writing assignments and among whom was the largest possible proportion of the target group. Teacher cooperation was another issue. The number of participants in our study was reduced somewhat because some of the faculty who teach Humanities 201

or History 201 were not comfortable participating. In spite of a concerted effort to achieve an optimum design, we still included a smaller number of writers in certain categories than we had hoped; in our final sample of 432 students (65 of the original 497 failed to complete all of the work), there were 182 sophomores, of whom 55 had AP credit. Nevertheless, our sample of students represented a broad cross-section of students from all colleges at BYU and the profile of sophomore students, in particular, matched the general profile of students accepted here. We considered recruiting teachers from other History of Civilization courses—e.g., in music or art—to participate, but we decided against that because we did not think we could devise uniform writing prompts that would work for various kinds of courses. Also, the complexities of gathering, photocopying, and evaluating many more essays, as well as training more raters, would have made the study nearly impossible to administer within a semester. Although we intended the writing tasks to be general enough to minimize structural differences among the humanities and history courses, we acknowledge some variability among the participating instructors in course content, approaches to teaching, and methods of weighting the writing tasks in the students' final grades.

Despite the sampling limitations under which the study was conducted, we believe several important conclusions can be drawn. First, the results indicate the value of the high school AP English experience. It is important to note here that we do not advocate abolishing AP English courses. We encourage students to take the most challenging courses available in their high school curriculum. But we believe that our results indicate as well the importance of students' receiving additional writing instruction in college. The students in our sample who performed best overall were those who combined a quality AP experience (as indicated by a 4 or 5 on the exam) with a first-year writing course. But students with only a 3 on the AP exam who had taken advanced FYC also performed better in this study than did their peers who bypassed the course. These results indicate that taking FYC, particularly advanced FYC, adds value to the AP experience.

Second, these results demonstrate that our study is more sensitive to distinctions in writing competence than the earlier studies used to deny a need to change BYU's AP policy. Those grade-based analyses found no differences between students with AP scores of 3, 4, and 5, nor did they find differences between students who had AP credit and those who didn't. As we have noted, because grades in writing courses may measure many factors besides ability to write and may vary widely from section to section, they constitute a very blunt instrument to use for assessing writing proficiency.

Third, our results show that students who score a 3 on the AP exam and do not take a first-year writing course are likely to suffer real consequences in sophomore courses that require writing assignments. Although differences between AP and non-AP students may diminish by the senior year, we maintain that this does not constitute a good reason to excuse students with AP scores of 3 from taking FYC. We believe that students are likely to struggle, particularly in the freshman and sophomore years, if they bypass the experience of taking a college writing course at the institution where they have matriculated.⁴ At BYU, as is likely true of most institutions, FYC provides an introduction to the discourse of the university, to a university library, and to genres of writing students have likely not encountered before.

Fourth, we believe this study justifies raising the AP English bar at BYU from 3 to 4. We conclude this even though there was no significant difference between the mean scores of non-AP students who had completed English 115, our standard FYC course (5.13), and those who chose to bypass FYC because they received an AP score of 3 (5.42). From the perspective of the essay raters, however, the performance of neither group was satisfactory. Their papers were characterized as “limited or unevenly developed in their reasoning” and having “lapses in diction or syntax.” Their background preparation was apparently inadequate for the analytical writing this assignment required. Because we did find that value can be added to the AP experience when students take advanced FYC, we recommend that students with AP scores of 3 be required to take an advanced FYC course.

Fifth, we believe the results indicate that changes are needed in BYU’s English 115 course to boost all students to a higher level of performance. We were disappointed to see that the English 115 students did not perform as well as we expected; nevertheless, we believe they still benefit from the course in ways noted earlier and in other ways we will explain in a subsequent article based on the survey data we collected. An analysis of circumstances that likely affected the performance of English 115 students relative to that of AP students who took FYC may help clarify why we still believe the bar should be raised from 3 to 4. One possible reason for the difference between scores in English 115 and in advanced FYC is that students self-select the advanced FYC courses; they are very likely those who see themselves as already good writers, yet they are motivated to experience a challenge. Another possible reason is that there are curricular differences between English 115 and advanced FYC that may have led to different performance levels. The advanced FYC courses focus more on inquiry, argumentation, and critical analysis, and require more sustained and in-depth research. For example, in the English 200 course, students research and

write within the same topic area all semester. This shared inquiry requires of students a sustained engagement with other writers in the class, as well as with the subject matter.

Yet another very likely reason is the differences in the teachers of English 115 and the advanced FYC courses. Each year more than 200 sections of English 115 are taught by a constantly changing corps of about sixty MA students, supervised by two full-time faculty. Although individual teachers in this group become very proficient, on the whole they are far less experienced than the relatively stable group of about thirty adjunct and full-time faculty who are assigned to teach more than fifty sections of advanced FYC each year. It may not be the wisest policy to have the English 115 students, who largely have not had the benefit of AP English in high school, taught by less experienced and generally less proficient teachers. We believe if we could improve the level of teaching in English 115 classes, the students would likely demonstrate higher levels of proficiency.

Some of our conclusions may be applicable to writing programs at other institutions. Because this study shows that value is added to the AP experience when students take an advanced first-year writing course, we recommend that policies allowing exemption based on AP scores should be discarded in favor of policies that are true to the name “advanced placement.” If programs offer only one approach to writing instruction—the monolithic first-year writing course—then it is more difficult to argue that this one course will meet the needs of those with an excellent AP experience in high school. Certainly, colleges can make a stronger case against giving away part of their curriculum to the high schools when they offer a curriculum that truly expands upon what students have learned in high school. Raising the bar means that colleges need to offer better classes and a wider array of classes, including alternatives to the standard first-year course.

As a result of this study, a formal proposal has been submitted to BYU’s administration to raise the bar from 3 to 4 before allowing students to bypass FYC. This proposal is a compromise because, ideally, all students should take some form of FYC. But because we believe students with AP scores of 3 stand to benefit the most from our courses and because they are the least likely to enroll in FYC on their own, we are most eager to see the requirement changed for their sake. Our proposal also takes into account pragmatic financial concerns. Because raising the bar will necessitate more sections of FYC, we have proposed to expand sections of our advanced FYC courses (English 200, Honors 200, and Philosophy 200) to enroll more students since the study shows these courses seem to be particularly effective in adding value to students’ writing. This expansion can be accomplished without adding significantly to staffing or space needs by scheduling the addi-

tional sections in the winter semester. Enrollments in FYC tend to be much higher in the fall than in the winter, and we can employ teachers and use space more evenly throughout the year by adding more sections of advanced FYC in the winter.

A second important feature of our proposal is to increase the number of full-time composition lecturers in the English Department from four to six. We propose that each of these six experienced teachers be given a one-course reduction in their teaching load to serve as a mentor to a group of six to eight master's degree students teaching English 115. Having these mentors spend ten hours a week observing and mentoring a small group of teachers will, we hope, improve the quality of teaching and learning that takes place in our English 115 classrooms. As of this writing, the findings of this study and the resulting proposal have been presented to important university committees, where they have been received positively. The academic vice-president has shown interest in the proposal, and we expect that funds will be sought from our Board of Trustees in the next academic year.

We offer this research to other universities that are also struggling to establish or revise their policy on what credit, if any, to grant for AP English scores. In the process of making AP policy, program administrators should certainly do their own research within the context of their own institution. In "Setting Your Credit and Placement Policy," the College Board provides several possibilities, including conducting a study such as the one reported here (see College Board, *AP Central*). We encourage other universities to conduct similar research so that more empirical data can be accumulated to add to or contradict our findings. We offer the design of our study as one that might be used or modified by other institutions for whom this issue is pertinent.⁵

NOTES

¹ "We" are a team of five professors and two graduate students from the departments of English, Instructional Psychology and Technology, and Molecular and Microbiology, as well as Freshman Academy, BYU's learning community program. Dr. William S. Bradshaw, a molecular biologist, deserves the credit for spearheading the research reported here because of his concerns about the quality of student writing.

² The university administration has been willing to advise and encourage students with AP credit to take first-year writing. The Fall 2003 class schedule tells students: "We strongly encourage you to take a First-Year Writing course even if you have received AP credit for English, because the skills taught in these courses are essential to a successful university education" (emphasis in original).

³ In the sample of students we studied, only four students had no AP credit and had not taken FYC. Because the number was so small, we determined not to compare this group to the other three groups.

⁴ The following two essay excerpts (written in response to the Miller prompt) represent the performance of (1) a student who scored a 3 on the AP exam and did not take FYC and (2) a student who had the same AP score and completed FYC. Readers should note that these examples were handpicked to represent the two ends of our holistic rating scale, and do not imply that all students of the same AP/FYC grouping will perform at exactly these levels.

Excerpt (1). Rater scores for this person were: Miller essay, 2, 2; Hart prompt 2, 4.

“Trials are a time to make a person stronger, it is a refiner’s fire, purifying of all the characters infirmities. Tragedy brings out either the best or the worst in us. We will either shun the fight or proudly go out to fight for our lives. When the true character is revealed and it is shown that the character wasn’t as strong as once believed, a sense of pessimism can emanate. But besides those few exceptions, an opposite sense of hope and optimism will be felt when the character will fight for their life and their personal dignity; and they come out the victor! By witnessing a tragedy transpire to a person on stage, observing the character have the courage to come out a conqueror can magnify a spirit of optimism. Which gives the audience a sense of hope that they can also be the vanquishers in their personal tragedies.”

Excerpt (2). Rater scores for this person were: Miller essay, 8, 9; Hart prompt 6, 6.

“Though optimism is an integral part of a tragedy, inherent also is the hopeless state of mankind’s endeavors. Tragedy forces us to realize that no matter how good a person is, he cannot save himself. Society may value his qualities, but they will not redeem him from his judgmental error. The Iliad and Second Samuel show how God, the gods, fate, or nature must be appeased—they cannot be ignored. Tragic heroes, obsessed with their own desires and personal dignity, fall short of their potential. Miller says that the hero ‘has fought a battle he could not possibly have won.’ The reader hopes for victory, but knows that it is impossible. Stories like those of King David and Achilles allow us to see that there is greatness in mankind, however they inevitably end with a frightening view of the instability of this greatness. Tragedy must educate, teaching optimism—to see the potential in one’s fellow man—and the futility of a man’s attempt to realize his own potential.”

⁵We gratefully acknowledge the cooperation and assistance of the following teachers: Jon Green, Jane Hinckley, Joseph Parry, Rita Wright, and Hal Black. We also thank the following graders, who spent several Saturdays and evenings reading essays: Martha Baumgarten, Samantha Butterworth, Julie Carter, Kristen Farnworth, Jennifer Gonzalez, Debbie Harrison, Lisa Johnson, Jennifer Nieves, Deirdre Paulsen, Angela Thompson, Sara Wagner, and Heidi Yates.

WORKS CITED

- Bodenhausen, Judith. Do public and private schools differ in the performance of their students on Advanced Placement tests? *Paper Presented at the Annual Meeting of the American Educational Research Association, San Francisco, March 27–31, 1989*. ERIC Document Reproduction Service, ED 312 279.
- Brennan, Robert L. *Generalizability Theory*. New York: Springer Verlag, 2001.
- Bridgeman, Brent and Charles Lewis. “The Relationship of Essay and Multiple-Choice Scores with Grades in College Courses.” *Journal of Educational Measurement* 31.1 (1994): 37–50.
- Burnham, Paul S. and Benjamin A. Hewitt. “Advanced Placement Scores: Their Predictive Validity.” *Educational and Psychological Measurement* 31.4 (1971): 939–45.
- Camara, Wayne, Neil J. Dorans, Rick Morgan, and Carol Myford. “Advanced Placement: Access, Not Exclusion.” *Education Policy Analysis Archives* 8 (2000). 1 August 2000. <<http://epaa.asu.edu/epaa/v8n40.html>>.
- Chamberlain, Philip C., et al. “Does Advanced Placement Continue Throughout the Undergraduate Years?” *College and University* 53 (1978): 195–200.
- College Board. *Access to Excellence: A Guide to the Advanced Placement Program*. Princeton: College Entrance Examination Board. (May 2000).
- AP Central*. <<http://apcentral.collegeboard.com>>. Retrieved July 9, 2003.
- AP Central*. <http://apcentral.collegeboard.com/colleges/setting_policy/0,3060,154-179-0-0,00.html>. Retrieved October 1, 2003.
- The 200 Colleges and Universities Receiving the Greatest Number of AP Grades*. Princeton: College Entrance Examination Board. (1999).
- Creech, Joseph D. “Challenging Students to Higher Standards through Advanced Placement.” (1995). ERIC Document Reproduction Service, ED 387 753.
- Cronbach, Lee J., G. C. Gleser, H. Nanda, and N. Rajaratman. *The Dependability of Behavioral Measurements: Theory of Generalizability for Scores and Profiles*. New York: Wiley, 1972.
- Cusker, Elaine Roberta. “A Study of the Use of Advanced Placement Credit by Undergraduates Entering Binghamton University in 1990 (New York).” Diss. State U of NY at Buffalo, 1999. UMI: AAT 9931474.
- Daly, John A. and M. D. Miller. “The Empirical Development of an Instrument to Measure Writing Apprehension. *Research in the Teaching of English* 9 (1975): 242–49.
- Dvorak, Jack. “Journalism Student Performance on Advanced Placement Exams.” *Journalism and Mass Communication Educator* 53.3 (1998): 4–12.
- Foster, David. “The Theory of AP English: A Critique.” Eds. Gary A. Olson, Elizabeth Metzger, and Evelyn Ashton–Jones. *Advanced Placement English: Theory, Politics and Pedagogy*. Portsmouth: Boynton Cook, 1989. 96–115.

- Ganeshanathan, Vasugi V. "AP Program Faces New Criticism Over Its Testing Standards." *The Chronicle of Higher Education* (14 July 2000): A45–46.
- Hart, Edward L. "The Need Beyond Reason." *Brigham Young University Studies* 16 (1976): 517-26.
- LeMy, M. "Academic Achievement of Students who receive AP Credit." *College and University* 60 (1985): 155–59.
- Lichten, William. "Whither Advanced Placement?" Education Policy Analysis Archives 8 (2000). 24 June 2000. <http://epaa.asu.edu/epaa/v8n29.html>
- Mahala, Daniel and Michael Vivion. "The Role of AP and the Composition Program." *WPA: Writing Program Administration*, 17.1–2 (1993): 43–56.
- Maxwell, Scott E. and Harold D. Delaney. *Designing Experiments and Analyzing Data*, 2nd ed. Mahwah, NJ: Erlbaum, 2004.
- Mercurio, Joseph A. "College Courses in the High School: A Four-Year Follow-Up of the Syracuse University Class of 1977." *College and University* 58 (1982): 5–18.
- Mercurio, Joseph A., et al. "College Credit Earned in High School: Comparing Student Performance in Project Advance and Advanced Placement." *College and University* 59 (1983): 74–86.
- Miller, Arthur. "Tragedy and the Common Man." *The Theater Essays of Arthur Miller*. Ed. Robert A. Martin. New York: Viking, 1978. 3–7.
- Modu, Christopher C. "The Validity of the Advanced Placement English Language and Composition Examination." *College English* 43.6 (1981), 609–620.
- Morgan, Rick and Len Ramist. *Advanced Placement Students in College: An Investigation of Course Grades at 21 Colleges*. Unpublished Statistical Report. Princeton: Educational Testing Services. February 1998. Report No. SR-98–13.
- Richardson, Mary W. *Educational Usefulness of the Advanced Placement English Course in New York State*. Diss. U of Sarasota, 1978. ERIC Document Reproduction Service, ED 163 037.
- Shell, Duane F., Carolyn C. Murphy, and Roger H. Bruning. "Self-efficacy and Outcome Expectancy Mechanisms in Reading and Writing Achievement." *Journal of Educational Psychology*, 81 (1989): 91–100.
- Shrout, Patrick E. and Joseph L. Fleiss. "Intraclass Correlations: Uses in Assessing Rater Reliability." *Psychological Bulletin*, 86 (1979): 420–428.
- Vopat, James B. "The Politics of Advanced Placement English." *Advanced Placement English: Theory, Politics, and Pedagogy*. [Ed.] Gary A. Olson, Elizabeth Metzger, and Evelyn Ashton–Jones. Portsmouth, NH: Boynton/Cook, 1989. 52–64.
- Willingham, Warren W. and Margaret Morris. *Four Years Later: A Longitudinal Study of Advanced Placement Students in College*. Princeton: Educational Testing Services. 1986. Report, No. ETS-RR-85-46. ERIC Document Reproduction Service, ED 280 358.

Table 1.
 Entering Freshmen at BYU: Percent Fulfilling Freshman Composition Requirement by Course Option in 2000-2001

	Percent enrolling in English 115	Percent enrolling in Advanced FYC	No composition course
Students with AP English credit	8.2	31.4	60.4
Students without AP English credit	68.6	19.9	11.5
All freshmen	47.3	23.9	28.7

Number of students with AP English credit = 1,383; students without AP English credit = 2,535; total size of freshman class = 3,918.

Table 2.
 Essay Scores by AP Status (Sophomores)

Group	N	Mean	Std. error
AP credit + FYC	22	6.47	0.24
AP credit, no FYC	33	5.42*	0.20
No AP credit, FYC	127	5.13**	0.11
Total	182	5.35	0.09

*Comparison of (AP, no FYC) with (AP + FYC) is significant, $p = 0.005$

**Comparison of (no AP, FYC) with (AP + FYC) is significant, $p < 0.001$
 (p-values derived from Games-Howell posthoc tests for samples of unequal size and unequal variances)

Source	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	33.47	2	16.74	11.99	.000
Within Groups	249.84	179	1.40		
Total	283.32	181			

Table 3. Essay Scores by AP Scores (Sophomores)

Group	AP score = 3	AP score > 3	Row Means
AP cr. + FYC	6.29 ± 0.28 n = 7	6.55 ± 0.34 n = 15	6.42* ± 0.24 n = 22
AP cr., no FYC	4.91 ± 0.24 n = 17	5.95 ± 0.26 n = 16	5.43* ± 0.20 n = 33
Column Means	5.60* ± 0.23 n = 24	6.25* ± 0.21 n = 31	5.84 ± 0.17 n = 55

Row and column means are unweighted, following Maxwell and Delaney.

Numbers reported are essay means ± std error of the mean.

*p < 0.05 interaction between cells: p = 0.215

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	23.80	3	7.94	6.85	0.001
Intercept	1697.82	1	1697.82	1465.16	0.000
AP category	5.15	1	5.15	4.45	0.04
First-year Composition	11.74	1	11.74	10.13	0.002
AP category by FYC	1.83	1	1.83	1.58	0.215
Error	59.10	51	1.16		
Total	1956.38	55			
Corrected Total	82.90	54			

Table 4. Freshman Composition Courses Completed by “Best” and “Poorest” Writers in Our Sample

	English 115 (%)	Advanced FYC (%)	None (%)	# of students
“Poorest writers”	63.3	12.2	24.4	90
“Best writers”	33.3	37.6	29.2	72
All students	51.2	22.2	27.2	162

Chi-square = 15.49, df = 2, p < 0.0004

Table 5. Essay scores by hours in residence (n = 429)

Hours in residence	# of students	Mean essay score ± std. error
0–29.9	161	5.4 ± 0.12
30–59.9	171	5.5 ± 0.10
60–89.9	70	5.5 ± 0.16
90+	27	5.3 ± 0.33

Univariate ANOVA: no significant differences between groups, p = 0.781

APPENDIX A

The Task: Read the attached essay by Arthur Miller entitled “Tragedy and the Common Man.” Miller asserts that “tragedy implies more optimism in its author than does comedy . . . because the possibility of victory is inherent in tragedy.” Are you persuaded that Miller’s notion is true, based on the arguments in his essay? You are to write an essay in which you agree with, disagree with, or modify Miller’s proposition. In your essay, you must do at least the following two things: First, demonstrate that you understand Miller’s claims by summing up his argument in your own words. Second, assert and support your own thesis about the “tragedy-optimism connection” through an analysis of a tragic figure you’ve encountered up to this point in the course (from the Bible or in Greek drama or culture). Please note that your essay should be unified, i.e., it should both demonstrate your ability to read and restate Miller’s argument effectively while at the same time advancing your own thesis and argument in the context of a literary tragic figure.

Essay Requirement: Assume that your audience consists of Dr. _____, your classmates, and any other educated persons—a community of learners engaged in a study of world civilization. Your essay should be three pages in length, double-spaced, with standard one-inch margins. Please word-process rather than hand-write your essay, using a 12-point font.

Due Date: Please turn in your paper at the beginning of class one week from today, _____, Sept. _____. The next page in this packet is a cover sheet to be used as the first page of your essay. Please record your name and social security number on the cover sheet, but do not write your name on any of the three pages of the essay itself. This paper will count up to _____ points toward your grade in the course.

The Expectation: This assignment allows you to demonstrate that you can read carefully and express important ideas through writing that is clear and interesting. Please follow a process that will result in your very best effort.

APPENDIX B

he Task: Read the attached essay by Edward Hart entitled “The Need Beyond Reason.” As you read, think about Hart’s argument with respect to your study of history, art, architecture, literature, and/or music in your Civ 201 course thus far in the semester. Then write an essay in which you agree with, disagree with, or modify Hart’s central assertion. In your essay, you must do at least the following two things: First, demonstrate that you understand Hart’s claims by summing up his argument in your own words. Second, assert and support your own thesis about the value of studying the humanities. Please note that your essay should be unified, i.e., it should both demonstrate your ability to read and restate Hart’s argument effectively while at the same time advancing your own thesis and argument.

Essay Requirement: Assume that your audience will include not only Prof. _____ and your classmates, but also any educated person interested in the question of the value of humanistic study. Your essay should be three pages in length, double-spaced, with standard one-inch margins. Please word-process rather than hand-write your essay, using a 12-point font.

Due Date: Please turn in your paper at the beginning of class one week from today, _____, Oct._____. The next page in this packet is a cover sheet to be used as the first page of your essay. Please record your name and social security number on the cover sheet, but do not write your name on any of the three pages of the essay itself. This paper will count up to _____ points toward your grade in the course.

The Expectation: This assignment allows you to demonstrate that you can read carefully and express important ideas through writing that is clear and interesting. Please follow a process that will result in your very best effort.

APPENDIX C: RUBRIC FOR MILLER ESSAY

SUPERIOR

9. Essays earning the score of 9 meet all the criteria for 8 papers and, in addition, are particularly well-reasoned or demonstrate particular stylistic command.

8. Essays earning a score of 8 effectively engage Miller's claim that tragedy is more optimistic than comedy by analyzing, in an integrated way, the applicability of Miller's claim to the tragic figure(s) the student has chosen.

ADEQUATE

7. Essays earning the score of 7 fit the description of 6 essays but are distinguished from them by fuller analysis or stronger prose style.

6. Essays earning the score of 6 adequately respond to the prompt. They represent Miller's claim accurately and state a position in response (agree, disagree, modify). They illustrate their argument through a discussion of a tragic figure(s), but in a more mechanical or less integrated fashion than a more effective essay. A few lapses in diction or syntax may be present, but for the most part the prose of 6 essays conveys their writers' ideas clearly.

LIMITED

5. Essays earning the score of 5 respond adequately to the prompt, but their own reasoning is limited or unevenly developed. They may address Miller's claim or discuss a tragic figure, but in an unbalanced manner. A few lapses in diction or syntax may be present, but for the most part the prose of 5 essays conveys their writers' ideas clearly.

INADEQUATE

4. Essays earning the score of 4 respond inadequately to the question's tasks. They may misrepresent Miller's claim or misapply Miller's essay to a tragic figure(s). They may also discuss Miller's claim and tragedy without making a direct connection between the two. The prose of 4 essays usually conveys the writers' ideas adequately, but may suggest inconsistent control over such elements of writing as organization, diction, and syntax. These essays may read more like a first draft than a finished draft.

3. Essays earning the score of 3 are described by the criteria for the score of 4 but are less perceptive in their reading of the prompt, less successful in analysis, or are particularly inconsistent in their control of the elements of writing.

LITTLE OR NO SUCCESS

2. Essays earning the score of 2 demonstrate little or no success in responding to the prompt. Some may substitute another task, such as merely summarizing Miller's essay or talking generally about tragedy. These essays may indicate a simplistic understanding of the writing task with very little success in analysis. The prose of 2 papers may reveal consistent weaknesses in grammar or another of the basic elements of composition.

1. Essays earning the score of 1 are described by the criteria for the score of 2, but are particularly simplistic in their responses to the prompt, demonstrate little understanding of analysis, or are particularly weak in their control of language.

0. Essays earning the score of 0 are completely off topic or are unreadable.

Appendix D: Rubric for Hart Essay

SUPERIOR

9. Essays earning the score of 9 meet all the criteria for 8 papers and, in addition, are particularly well-reasoned *or* demonstrate particular stylistic command.

8. Essays earning a score of 8 effectively engage, in an integrated manner, Hart's claim about the value of studying the humanities, by asserting and supporting a thesis that agrees with, disagrees with, or modifies Hart's claim, drawing on the student's understanding of history, art, architecture, literature, and/or music.

ADEQUATE

7. Essays earning the score of 7 fit the description of 6 essays but are distinguished from them by fuller analysis *or* stronger prose style.

6. Essays earning the score of 6 adequately respond to the prompt. They represent Hart's claim accurately and state a position in response (agree, disagree, modify). They illustrate their argument through a discussion of the humanities, but in a more mechanical, more general, or less integrated

fashion than a more effective essay. A few lapses in diction or syntax may be present, but for the most part the prose of 6 essays conveys their writers' ideas clearly.

LIMITED

5. Essays earning the score of 5 respond adequately to the prompt, but their own reasoning is limited or unevenly developed. They may address Hart's claim or discuss the value of the humanities generally, but in an unbalanced manner. A few lapses in diction or syntax may be present, but for the most part the prose of 5 essays conveys their writers' ideas clearly.

INADEQUATE

4. Essays earning the score of 4 respond inadequately to the question's tasks. They may misrepresent Hart's claim or misapply Hart's essay to examples from the student's study of the humanities, or may discuss Hart's essay and the humanities without making an explicit logical connection between the two. The prose of 4 essays usually conveys the writers' ideas adequately, but may suggest inconsistent control over such elements of writing as organization, diction, and syntax. These essays may read more like a first draft than a finished draft.

3. Essays earning the score of 3 are described by the criteria for the score of 4 but are less perceptive in their reading of the prompt, less successful in analysis, or are particularly inconsistent in their control of the elements of writing.

LITTLE OR NO SUCCESS

2. Essays earning the score of 2 demonstrate little or no success in responding to the prompt. Some may substitute another task, such as merely summarizing Hart's essay or talking in very general terms about the humanities. These essays may indicate a simplistic understanding of the writing task with very little success in analysis. The prose of 2 papers may reveal consistent weaknesses in grammar or another of the basic elements of composition.

1. Essays earning the score of 1 are described by the criteria for the score of 2, but are particularly simplistic in their responses to the prompt, demonstrate little understanding of analysis, or are particularly weak in their control of language.