



---

*Research  
Report*

**Impact of Group Work and  
Extended Essay Writing on  
Online Advanced  
Placement Program®  
Student Performance**

**Phil Handwerk**

**Impact of Group Work and Extended Essay Writing on Online Advanced Placement  
Program<sup>®</sup> Student Performance**

Phil Handwerk  
ETS, Princeton, NJ

May 2007

As part of its educational and social mission and in fulfilling the organization's nonprofit charter and bylaws, ETS has and continues to learn from and also to lead research that furthers educational and measurement research to advance quality and equity in education and assessment for all users of the organization's products and services.

ETS Research Reports provide preliminary and limited dissemination of ETS research prior to publication. To obtain a PDF or a print copy of a report, please visit:

<http://www.ets.org/research/contact.html>

Copyright © 2007 by Educational Testing Service. All rights reserved.

ETS and the ETS logo are registered trademarks of Educational Testing Service (ETS). The ADVANCED PLACEMENT PROGRAM, AP, and PSAT/NMSQT are trademarks of the College Board.



## **Abstract**

Online high schools are growing significantly in number, popularity, and function. However, little empirical data has been published about the effectiveness of these institutions. This research examined the frequency of group work and extended essay writing among online Advanced Placement Program<sup>®</sup> (AP<sup>®</sup>) students, and how these tasks may have impacted their AP exam performance. Hierarchical, set-wise regression models found little to no impact of group work or extended essay writing on multiple-choice, free response, or composite scores across 9 AP exams administered from 2003–2005. These models included student sex, race, and year in school as independent variables; covariates were the number of previous AP exams taken by the student and PSAT/NMSQT<sup>®</sup> scores. Implications for online providers and for future research of online high schools are discussed.

Key words: AP, online, group work, essay writing

## Table of Contents

	Page
Introduction.....	1
Online High Schools.....	1
Group Work.....	3
Writing Extended Essays.....	4
AP and Online Classrooms, Group Work, and Extended Essay Writing.....	5
Research Questions.....	6
Methodology.....	7
Population and Sample.....	7
Background Questions.....	7
Dependent and Independent Variables.....	9
Research Design.....	14
Power and Sample Size.....	17
Results.....	17
Online-Only Students.....	17
Frequency and Impact of Group Work.....	20
Frequency and Impact of Extended Essay Writing.....	23
Discussion.....	24
Limitations of the Study.....	24
Online AP Students.....	25
How Frequently Are Online-Only AP Students Working in Pairs or Small Groups and Does It Impact Their AP Exam Performance?.....	26
How Frequently Are Online-Only AP Students Writing Extended Essays and Does It Impact Their AP Exam Performance?.....	28
Implications.....	29
For Online-Instruction Providers.....	29
For Future Research.....	30
References.....	30
Notes.....	36
Appendixes	

A - NAEP Tables on Group Work and Extended Essay Writing .....	37
B - AP Tables for Online Instruction.....	48
C - PSAT/NMSQT Tables on Group Work.....	55

## List of Tables

	Page
Table 1. AP Exam Counts for Online-Only Students .....	9
Table 2. Demographic Independent Variables for Initial Entry Into Regression Models .....	10
Table 3. Academic Performance/Experience Independent Variables and Covariates for Initial Entry Into Regression Models .....	13
Table 4. Independent Variables of Primary Interest .....	14
Table 5. Regression Models Summary by Independent Variables of Primary Interest.....	15
Table 6. Number of AP Exams Administered 2003–2005 by Degree of AP Online Instruction Across All AP Exams .....	18
Table 7. Frequency of Group Work by AP Exam for Online-Only Students.....	20
Table 8. Frequency of Writing Extended Essays by AP Exam for Online-Only Students.....	23

## **Introduction**

The remarkable growth of the Internet over the last several years has impacted not only corporations and homes but schools as well. More and more frequently, educators, students, and legislatures have been turning to technology as a means of providing information and instruction. Online high schools, or “state approved and/or regionally accredited school[s] that offers secondary credit courses through distance learning methods that include Internet-based delivery” (Clark, 2000, p. i), funded by public and private dollars, are growing significantly in number, popularity, and function.

So far, little evaluative information is available on these relatively new online high schools. There is a need for more research that will support “effective implementation and use of online learning in K-12 learning communities” (Blomeyer, 2002, p. 10). Vrasidas, Zembylas, and Chamberlain (2003) stated that, “studies are of critical importance for establishing a model for the development, delivery, support, and evaluation of distance education and online programs” (p. 201). Even more specifically, Marcel (2003) called for College Board® sponsorship of “studies leading to the definition of standards and guidelines for vendors and virtual schools to use when developing and delivering online [Advanced Placement Program® (AP®)] courses” (p. 15).

This section will discuss briefly the history of online high schools, online high schools’ relationship to AP, and why investigations into who is going online as well as the impact of group work and extended essay writing on online AP student performance is needed.

### ***Online High Schools***

Online high schools have been established to meet a variety of educational needs. Among these needs are supplementing existing schools’ curriculum; providing courses to the home- or hospital-bound student; providing courses to those who may not be thriving in the traditional classroom; and offering equitable access to advanced coursework for districts with limited means (often rural or urban schools).

Twenty-two states have already established state-sanctioned, state-level online schools. In addition, 16 states have approved online charter schools within their state. Only 19 states have neither a state-level online school nor an online charter school. It has been estimated that nearly 25% of all public K-12 schools currently offer some form of online instruction and nearly 300,000 high school students attend at least one class online (Wood, 2005).



Many foresee online high schools playing an even greater role in the future. According to a National School Boards Foundation survey, “28 percent of school leaders believe at least one in five of their students will receive a substantial portion of their instruction over the Internet in the next three years” (Vail, 2002). The National Association of State Boards of Education (NASBE, 2001) reported, “E-learning will improve American education in valuable ways and should be universally implemented as soon as possible” (p. 6). William Fitzsimmons, Harvard University's Dean of Admissions and Financial Aid, predicted, “Online classes will be a necessity for years to come because the gap between the haves and have-nots continues to widen” (Jan, 2005).

*Online AP courses.* Perhaps the most frequently requested and offered courses are those focusing on the College Board's AP.<sup>1</sup> Clark (2000) stated, “AP has been a driving force in the development of several state-level virtual high schools” (p. 19). For many, provision of online access to AP courses helps level the academic playing field by allowing students who otherwise would not have exposure to rigorous academic coursework in high school to experience that challenge online (Dillon, 2005; Education Week, 2005b; Jan, 2005).

Frequently, online high schools offering AP courses have relied on federal dollars to develop their institutions. The U.S. Department of Education, through its Advanced Placement Incentive Program (APIP), aids states in making AP exams available to low-income individuals. APIP also supports other initiatives to increase AP enrollment, exam taking, and course availability for low-income students. Among the projects supported by APIP grants in the past have been those involving the creation of online AP courses (Clark, 2000).

*Online high school performance research.* Because many online high schools are only a few years old there has been very little empirical research reported on their impact (Marcel, 2003; Vail, 2001). As summarized by a public school district superintendent in Ohio, “This whole thing is moving so fast that it's caught everybody by surprise” (Cook, 2002, p. 1).

Although not monumental in volume, some assessment information has been found regarding performance of online high schools (Bigbie & McCarroll, 2000; Chen, Elbaum, & Walsh, 1999; Clark, 2001; Doherty, 2002; Interactive Educational Systems Design, Inc., 2002; National School Boards Foundation, 2002; University of California College Preparatory Initiative, 2002). These studies attempted to answer the question of how online high schools are succeeding in one of three ways: looking at the enrollment figures for online high schools; teacher, student, and parent satisfaction surveys; and performance on national or local achievement tests,

completion rates, and other measures of quality as a way to evaluate the institution. The only consistent finding from this literature is that the enrollments are growing across online schools. Gaps in the research exist with regard to who are these online students, and whether or not they are similar to traditional students.

While little research has been found that analyzes the success of online high schools, much has been written regarding what experts believe should be taken into account when creating online courses. Consistent themes are that the courses ought to make use of a variety of technological techniques, be interactive, and emphasize student-to-student and student-to-teacher communication (Bauer & Anderson, 2000; Lopez, 2003). Student-centered, constructivist, online courses are often recommended where the instructor frames the discussions to be handled through student-to-student and/or student-to-teacher communication.

*Criticisms of online education.* The advent and growth of online classrooms has not been received without its share of criticism. One of the concerns returns to the issue of little empirical research being available on the efficacy of online high schools (Blomeyer, 2002). Since so few studies or evidence are available, much of what is believed to be effective is based on opinion or anecdotes (Education Week, 2005a; Marcel, 2003). Others list the large time commitment placed upon online teachers and students' underestimation of online workload as challenges remaining to successful online education (Vail, 2001; Wood, 2005). Detractors also warn that many online high schools are hiring successful classroom teachers and having them teach online, only to find they are not succeeding—a procedure not advocated by online experts themselves.

One of the most frequent concerns voiced regarding online high schools is a fear that the students will suffer isolation from their peers and teachers (Russell, 2001). Critics worry the result may be “socially isolated students who are taught by machines and lack human contact” (Vail, 2002).

### ***Group Work***

With so much apprehension regarding online students potentially feeling isolated, one important area in the study of online schools is the impact of group work among online students. Although most of the studies concerning group work in academia focus on traditional brick-and-mortar classes, a few reported the impact of group work among distance education students. Group work was found to have had a positive impact on online students' interpersonal skills, sense of

responsibility, understanding of material, and confidence (Baskin, 2001; Kear, 2004; Parker, 1996). Fisher, Thompson, and Silverberg (2005) reported that group work was helpful in overcoming the sense of isolation among online students.

In a study comparing the impact of group work on online versus traditional undergraduate students, Baskin (2001) found the traditional students reporting higher levels of satisfaction with the group work experience compared to the online students. Still, Baskin did report benefits to the online students from working in groups. He cautioned, “The issue here is not really about ‘which’ [online or traditional group work] is better, but about how group work epistemology can be enhanced through the adaptation and integration of new learning technologies” (p. 44).

Additional research has been conducted on the frequency of group work among traditional high schools students. These studies tended to focus on how working in groups impacts students’ academic performance, an angle on the group work question not yet explored in online education. Over the years, numerous questions regarding group work have appeared on the National Assessment of Education Progress (NAEP) assessments. Generally, 12<sup>th</sup> grade students who engage in group work tended to perform better on the various NAEP assessments than did those who indicated they never work in groups, although there is no compelling evidence that more frequent group work necessarily improves scores (see Tables A1–A7). Another consistent finding from the NAEP data is that a substantial proportion of students do not engage in group work. Similarly, Langer (2001) found, “several studies have indicated that such groupings are not pervasive in American [high] schools” (p. 843).

### ***Writing Extended Essays***

In online education a great deal or perhaps all of the communication that occurs is written communication. Asynchronous communication exists where the instructor posts or emails an assignment to students, who in turn write their replies. Another example would be the use of electronic discussion boards where issues are raised and responses challenged, defended, and/or expanded one posting at a time over a period of hours, days, or weeks. Synchronous communication can occur by use of online chat with teachers and students logged into the same chat room in cyberspace and conversing in real time with one another via their keyboards.

Numerous benefits of writing have been reported in educational literature. Among these benefits is improved academic performance in general, and in reading scores in particular

(Bottoms & Bearman, 2000; Langer, 2001). Writing may also lead to gains in communication and higher order thinking skills (Angelo, 1995; Bottoms & Bearman, 2000; Marzano, 1993). Through the reinforcement of learning by writing, students have been reported to gain not only a better understanding of the content, but a better understanding of their own processes of learning (Cooper, 2004; Luria, 1971; Schraw & Dennison, 1994). To benefit most from writing experiences, several researchers and writing experts have encouraged students to regularly write extended essays, rather than having them complete short-answer assignments (Breland, Bonner, & Kubota, 1995; Bottoms & Bearman, 2000; Connor-Greene & Murdoch, 2000; Langer, 2001; Saunders & Scialfa, 2003).

The NAEP assessments also queried students regarding their writing experiences (see Tables A8–A16). Students in English courses tended to write at length more frequently than students in other subjects, and there seems to be a positive correlation between the frequency of writing and performance on the reading and writing assessments. Students in social science courses participated less frequently in writing activities of length, yet there appeared to be a positive relationship between the students who engaged in this sort of writing and assessment performance. Students in math and science courses wrote at length the least frequently. While there were some signs of positive correlation between writing and performance on some science questions, the relationship in math classes appeared negative. In other words, as frequency of writing among math students decreased, their assessment scores increased.

Researchers have found that including writing as part of academic coursework in courses other than English remains a popular and successful concept (Bottoms & Bearman, 2000; Connor-Greene & Murdoch, 2000; Marzano, 1993). Scott (2002) abridged the philosophy as “all educators are responsible for teaching reading, writing, and math, and their individual curriculum is a vehicle through which these basic skills are reinforced”

### ***AP and Online Classrooms, Group Work, and Extended Essay Writing***

The College Board has been particularly interested in all three topics discussed above with regard to AP. That is, they are actively gathering data focusing on AP students taking online courses, working in groups, and/or writing extended essays. This information is often used by the College Board and the AP Development Committees to better understand their AP students and the relationships between student characteristics/training and exam taking and performance. However,

much of this data has been analyzed separately and disaggregated across specific AP subjects. This report will provide a more comprehensive and integrated examination of these issues aggregated over three years and across nine subjects.

### **Research Questions**

This study looks at online students in the hope of identifying practices that may improve their academic performance. It is not the focus of this study to compare how the online students perform relative to traditional, brick-and-mortar students. Although online high schools are one of the newest forms of distance education, research on the effectiveness of distance education itself has been ongoing since the 1960s (Lockee, Moore, & Burton, 2001). During that time many studies were conducted comparing traditional classroom instruction to various modes of distance education. However, instead of comparison studies, researchers have suggested more detailed evaluation studies of the distance education environment. Comparison studies fail “to consider the many variables that work together to create an effective instructional experience” (Lockee, et al., 2001, pp. 60–61). By focusing on the mode of instruction, the research fails to investigate individual learning styles, learning theories, relationship of media and motivation, and the quality of interactions between students and teachers. Alternative lines of inquiry suggested include: determining if online students learned what the course was designed to teach; conducting quantitative and qualitative evaluation of interactions, examining students’ verbal and written behaviors; and examining how various aspects of media presentation impact learning.

The primary purpose of this study is to add to the current body of knowledge regarding who is taking online courses and what is successful in online classrooms. More specifically, this study provides a rare description of the population of students currently taking AP courses online. Additionally, the study will examine whether or not group work and/or writing extended essays impacts the academic performance of AP students who are attending online high schools.

To address the purposes of this study a number of research questions were considered. The questions focused on student background and demographics as well as two aspects of the online classroom. The research questions investigated were as follows:

1. Which students are taking online AP courses?
2. How frequently are online-only AP students working in pairs or small groups and does it impact their AP exam performance?

3. How frequently are online-only AP students writing extended essays and does it impact their AP exam performance?

## **Methodology**

### ***Population and Sample***

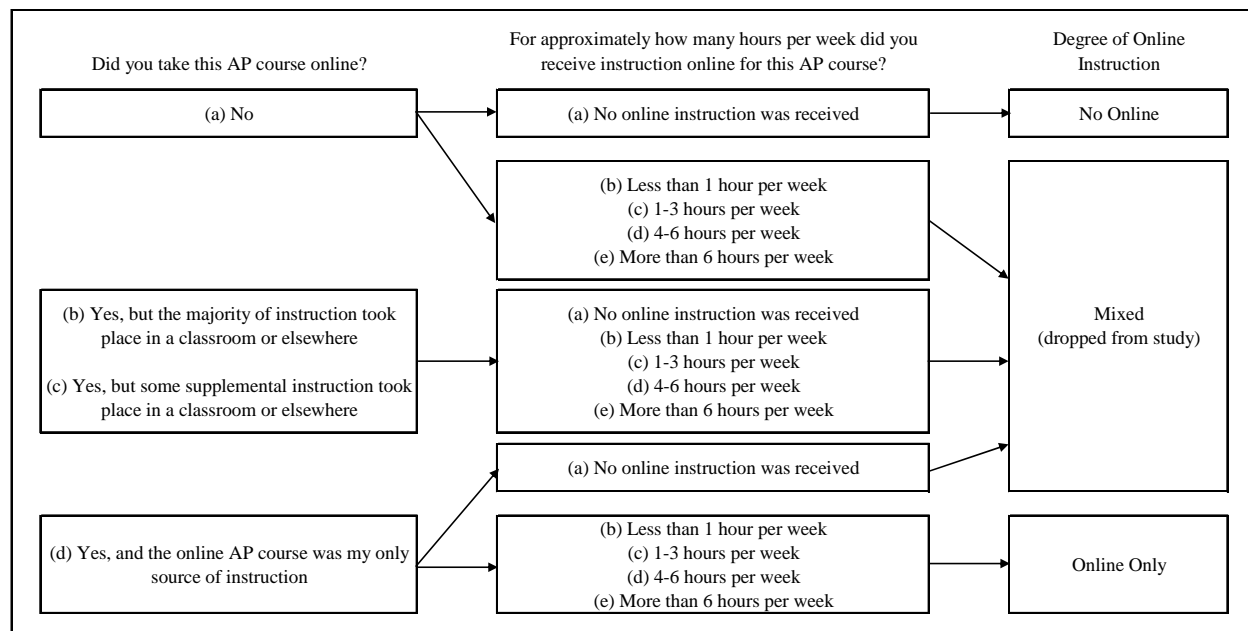
The population of interest is high school students who took an AP course online in the years 2003-2005. However, the College Board does not track all students who take an AP course. The College Board is only able to report the number of students who take an AP exam. Students who take an AP course may opt not to take the exam at the end of the year. College Board estimates that in 2005, 74% of the students who took an AP course sat for the exam in that subject (Cavanagh, 2006). Conversely, students may take an AP exam without ever formally enrolling in a course for that subject (e.g., through independent study), although such occurrences are presumed to be rare.

### ***Background Questions***

Prior to beginning each AP exam, the student is asked several background questions about themselves, their preparation for that exam, and their educational plans. Since 2003, students have answered two questions regarding taking their AP course online. The first question asked is, “Did you take this AP course online?” The possible responses are: (a) No; (b) Yes, but the majority of instruction took place in a classroom or elsewhere; (c) Yes, but some supplemental instruction took place in a classroom or elsewhere; or (d) Yes, and the online AP course was my only source of instruction. The second question asked is, “For approximately how many hours per week did you receive instruction online for this AP course?” Here the possible responses are: (a) No online instruction was received; (b) Less than 1 hour per week; (c) 1-3 hours per week; (d) 4-6 hours per week; or (e) More than 6 hours per week.

To be part of this study's sample classified as online-only the student on that particular exam must have answered (d) to the first question (“Yes, and the online AP course was my only source of instruction”) and (b) (c), (d), or (e) to the second question. To be classified as “no online,” the student must have indicated that he or she did not take the course online (first question, response “a”) and confirmed that “no online instruction was received” (second question

response “a”). All other students were coded as mixed because it is unclear as to how much, if any, instruction was received online and were dropped from the study. (See Figure 1.)



**Figure 1. Classification of degree of online instruction.**

Another background question asked of all examinees focuses on the use of group work in the AP class. For several years, the College Board has asked, “How often did you work in pairs or small groups for your AP class?” The response categories for the question are: (a) Never, (b) A few times a year, (c) Once or twice a month, (d) Once or twice a week, and (e) About every day.

Like group work, extended essay writing has been the focus of a background question asked of all AP examinees. The College Board has asked each student prior to beginning a particular exam, “How often did you write an extended essay (500 words or more) for your AP class?” The response categories for this question are: (a) Never, (b) A few times a year, (c) Once or twice a month, (d) Once or twice a week, and (e) About every day.

The overwhelming majority of all AP examinees (90% +) in 2003–2005 responded to each of these four background questions. With response rates typically near 100% for the online-only students, we can feel confident relying on responses to these background questions for the remainder of the analyses.

Although the College Board offers over 30 exams, this study only reviewed the impact of group work and extended writing using data from nine exams. These nine exams have over 150 students in each year (2003–2005) who indicated that their only source of instruction for that course was online and that they received at least some online instruction each week in that course. The exams studied were U.S. History, Microeconomics, Macroeconomics, English Language and Composition, English Literature and Composition, U.S. Government and Politics, Calculus AB, Psychology, and Statistics. The final group of online-only students presented in Table 1 is broken down by AP exam and also by year.

**Table 1**  
*AP Exam Counts for Online-Only Students*

AP exam	2003 # exams	2004 # exams	2005 # exams	2003–2005 combined # exams
U.S. History	367	352	370	1,089
Microeconomics	306	250	316	872
Macroeconomics	288	374	296	958
English Language & Composition	234	328	288	850
English Literature & Composition	265	317	353	935
U.S. Government & Politics	384	430	471	1,285
Calculus AB	188	186	181	555
Psychology	320	385	564	1,269
Statistics	158	191	175	524

### *Dependent and Independent Variables*

To determine the impact of group work and extended essay writing on academic performance in AP courses, a number of independent multiple regressions were performed (see following section for more detail). In these regressions, the dependent variable, academic performance, was measured by three scores from AP exams: the weighted multiple-choice section score, the weighted free response section score, and the combined or composite score.<sup>2</sup> The composite score is the sum of the weighted multiple-choice and weighted free response section scores. Using all three scores allows for investigation into whether or not the specific activities (i.e., group work or extended essay writing) have an impact on one particular section of the exam



or on the entire exam score. For example, one might presume writing extended essays to more likely affect the free response scores than the multiple-choice scores.

Previous research has found a number of variables related to the performance of students on the AP exams. Many of these variables were available in the AP datasets and were entered into the regression models as independent variables or covariates to account for the variation in student performance before the critical independent variables were included. These independent variables and covariates can be considered to fall into one of three categories: demographics, previous academic performance/experience, and pedagogical descriptors.

First, this report will examine the demographic independent variables included in the study. The College Board has reported differences in performance on AP exams by sex, race, and year in school (College Board, 2007a). Hence, these demographic variables were added in the regression models. Dummy variables were created for each of these demographics. For gender, the FEMALE variable was set to 1 when the student was female and 0 when the student was male. For ethnicity, four dummy variables were created (AF\_AM, ASIAN, HISP, and OTHR). If the student indicated their race to be white, the four dummy variables would all be set to 0. Similar dummy variables were created to indicate the student’s year in school, with seniors being the reference group (see Table 2).

**Table 2**

***Demographic Independent Variables for Initial Entry Into Regression Models***

Variable	Description
FEMALE (I)	Female (1) or male (0)
AF_AM (I)	African American (1) or not (0)
ASIAN (I)	Asian (1) or not (0)
HISP (I)	Hispanic (1) or not (0)
OTHR (I)	Other race (1) or not (0)
FR (I)	Freshman in high school (1) or not
SO (I)	Sophomore in high school (1) or not (0)
JR (I)	Junior in high school (1) or not (0)

*Note.* I = independent variable.

Consideration was also given to the students' previous academic performance and experience as independent variables or covariates to include in the regression models. A student's PSAT/NMSQT score has been found to correlate significantly with AP performance (Camara & Millsap, 1998). Therefore, it was decided that PSAT/NMSQT scores would be included in the regression models to account for some of the variation in online-only student performance. Online-only students' AP records were matched with PSAT/NMSQT data to identify students who have taken PSAT/NMSQTs and their performance on that assessment. The pairing of online-only AP students with PSAT/NMSQT scores yielded matches for 81% of the AP records. However, the characteristics of students who had a PSAT/NMSQT score and those who did not differed. Students without a PSAT/NMSQT score tended to be younger (freshmen or sophomores) and were more likely to be African American and less likely to be Asian. Also students with PSAT/NMSQT scores tended to score higher on their AP exams. To deal with these differences, one consideration was to separate online-only students for whom PSAT/NMSQT scores were found from those who did not have PSAT/NMSQT scores. Models could then be created separately for both groups of students and results compared. But in the world of the online classroom, students from schools where PSAT/NMSQTs are offered may be in the same online class with students where PSAT/NMSQTs are not. The ultimate audience for this research will be online high school administrators who, it is hoped, will find the results reported useful to course design. If separate models were presented, one for students with PSAT/NMSQT scores and one for students without, the information would be of less practical use to this audience, for they do not, presumably, have influence as to whether or not their students take the PSAT/NMSQT. Therefore, it was decided that to best serve the intended audience, models would be presented wherein all students, regardless of PSAT/NMSQT taking, would be analyzed together. To account for the differences found when looking at matching of PSAT/NMSQT scores with online-only AP students, a PSAT/NMSQT flag was created (PSAT2). If a PSAT/NMSQT score was found the variable PSAT2 was set to 1; if not, the variable equaled 0.

The PSAT/NMSQT flag provided a means of indicating whether or not a student's PSAT/NMSQT score was found, and thereby help account for some of differences in AP performance described previously. Still, it was desired to also include in the regression models PSAT/NMSQT scores as well as the PSAT/NMSQT flag to better account for the student's

academic ability. The PSAT/NMSQT scores (PSATV\_M) included in the study were the combined verbal (V) and math (M) score for that student.

The decision was made that for students for whom no PSAT/NMSQT score was found a score would be created via multiple imputation. Multiple imputation methods that ascribe more than one value for a missing item using available non-missing data have been found to create valid estimates of the missing data (Little & Rubin, 1987). Statistical programs were run that imputed five PSAT/NMSQT scores for each student for whom no PSAT/NMSQT score was found. A simple average was calculated from those five imputed PSAT/NMSQT scores and entered into that student's record for inclusion in the regression models. (See Table 3.)

While using multiple imputation techniques has proven effective in previous studies, using it to impute PSAT/NMSQT values where considerable portions of the data were missing may not have been appropriate. Across the nine exams, 81% of the records were found to have PSAT/NMSQT scores. Yet, English Language and Composition, English Literature and Composition, and Psychology each had over 20% of their students without PSAT/NMSQT scores. Relying on imputed data for over 20% of the records was deemed as potentially providing unreliable PSAT/NMSQT scores. It was therefore determined that another set of models would be built wherein only the PSAT/NMSQT flag, not the actual or imputed PSAT/NMSQT score, was included. By having two sets of models, one with PSAT/NMSQT scores and one without, for each of the nine exams, results could be compared to determine if there were consistencies or differences when the PSAT/NMSQT score was included and when it was not.

It was also desired to have the regression models account for whatever impact there might be on exam performance given AP experience. Adelman's 2006 report *The Toolbox Revisited* included the number of AP courses taken by a student during their academic career as an indicator of the student's academic rigor, finding that academic rigor impacted future academic performance. The decision was made to include the number of AP exams taken that year or previously as a related proxy for academic rigor and was included as a covariate in the models. Similarly, if the same student appeared more than once on the online-only listing (i.e., they indicated they are taking more than one AP course online that year or they appear on a previous year as having met the online-only definition) the number of the online AP exams was also used as a covariate. (See Table 3.)

**Table 3*****Academic Performance/Experience Independent Variables and Covariates for Initial Entry Into Regression Models***

Variable	Description
PSAT2 (I)	PSAT/NMSQT flag indicating if PSAT/NMSQT score found (1) or not (0)
PSATV_M (C)	PSAT/NMSQT score (imputed for students without found PSAT/NMSQT scores)
EXAMS_LIFETIME (C)	Total number of AP exams taken up to and including that point in time.
ONLINE_LIFETIME (C)	Total number of AP exams taken in 03-05 where student was classified as online.

*Note.* I = independent variable, C = covariate.

The third category of independent variables and covariates, pedagogical variables, included a measure of the amount of online instruction a student received. Initially, the average hours of weekly online instruction question discussed previously was dummy-coded into four separate variables. However, it was determined that coding in such a manner created variables that acted as suppressors. A new dichotomous variable was created indicating whether the student received at least 1–3 hours per week of online instruction. If the student received less than 1 hour per week of instruction, this variable equaled 0. This new variable (HRSONLINE\_GE1) was found to positively correlate with the outcome variables and was included in the regression models.

The two independent variables of primary interest in this study were the frequency of group work and the frequency of writing extended essays and part of the pedagogical variables category. Both questions, as mentioned above, have five values. Before being entered into the regression models, each variable was recoded into four dummy variables. For example, the group work variable became GROUPS\_2\_YR where 1 indicated the student replied “a few times a year” and 0 indicated they did not reply “a few times a year.” GROUPS\_3\_MO was coded 1 when the student answered “once or twice a month” and 0 when they did not respond “once or twice a month.” Similar dummy variables were created for the options where students worked in groups “once or twice a week” or “about every day.” Students who reported “never” working in groups

were recognized by the other four dummy variables all equaling 0. (See Table 4.) The same procedures were used to create dummy variables for the frequency of extended essay writing background question. Here again, “never” was used as the reference group.

**Table 4**  
*Independent Variables of Primary Interest*

Background question	Variable	Description
How often did you work in pairs or small groups for your AP class?	GROUPS_2_YR	A few times a year (1) or not (0)
	GROUPS_3_MO	Once or twice a month (1) or not (0)
	GROUPS_4_WK	Once or twice a week (1) or not (0)
	GROUPS_5_DAY	About every day (1) or not (0)
How often did you write an extended essay (500 words or more) for your AP class?	ESSAY_2_YR	A few times a year (1) or not (0)
	ESSAY_3_MO	Once or twice a month (1) or not (0)
	ESSAY_4_WK	Once or twice a week (1) or not (0)
	ESSAY_5_DAY	About every day (1) or not (0)

***Research Design***

Since this study looked at three measures of performance across nine AP exams for two primary independent variables (group work and extended essay writing), the desire was to find a model that best fit across all the outcomes and exams for both group work and extended essay writing. The study was focused on a more global picture for group work than extended essay writing rather than concentrating on each individual combination of exam and score individually.

Regression models were built first to address the research question focusing on group work. Then an entirely separate set of models were built focusing on the extended essay writing question. This yielded 54 regression models for the group work question (27 with PSAT/NMSQT flag only and 27 with PSAT/NMSQT flag and PSAT/NMSQT score) and 54 regression models for the extended essay writing question. (See Table 5.)

**Table 5*****Regression Models Summary by Independent Variables of Primary Interest***

Independent variables	Independent variables & covariates	PSAT/NMSQT variables & covariate	# of AP exams	AP scores	Total # of models
Group work	Demog, acad perf/exp, hrs online instruction	PSAT/NMSQT flag	9	* 3 scores (MC, FR, COMP)	27
	Demog, acad perf/exp, hrs online instruction	PSAT/NMSQT flag + PSAT/NMSQT score	9	* 3 scores (MC, FR, COMP)	27
Extended essay writing	Demog, acad perf/exp, hrs online instruction	PSAT/NMSQT flag	9	* 3 scores (MC, FR, COMP)	27
	Demog, acad perf/exp, hrs online instruction	PSAT/NMSQT flag + PSAT/NMSQT score	9	* 3 scores (MC, FR, COMP)	27

Hierarchical set-wise regression was used in this study since this technique allowed the main independent variables of interest to be entered into the models after all other independent variables and covariates have already been entered. This type of regression assumes there is an order, or hierarchy, to which the independent variables should be put into the model and that the variables can be put into the model in sets (hence, set-wise). The objective was to account for as much variance in the outcome measures as possible with the independent variables and covariates that were not of primary interest first. After that variance has been accounted for, the independent variables of primary interest would be entered. Here the demographic variables, academic performance/experience variables, and hours of online instruction were entered into the model initially. The group work or essay writing variables, respectively, were entered last into the models after effects of the other independent variables had been controlled for. Every model was also checked to be sure there were no violations of the normality of the residual distributions.

Multivariate multiple regression was initially considered as a means by which to create these models. The advantage of such analysis would be it allows for examination of more than one dependent variable simultaneously. However, concerns about multicollinearity among the section scores and the composite score ruled out using these procedures. Therefore, independent multiple

regression analyses were conducted on each of the three dependent variables, first with the group work dummy variables as the primary independent variables of interest, then with the essay writing dummy variable as the primary independent variables of interest.

Often students in evaluations are treated as independent of one another when in fact they do share a commonality (e.g., the same classroom in a school). This treatment violates an assumption of many statistical procedures, that being independence of observations. Some of the online-only students in this study, while not sitting in the same classroom, may have shared the same online provider and even the same online teacher. Hierarchical linear models (HLM) allow for the examination of the variance due to the individual student as well as that due to the shared classrooms (Raudenbush & Bryk, 2002).

However, the data available did not allow estimation of an HLM model since the second level cluster (i.e., online teacher) could not be clearly defined. Like traditional brick-and-mortar schools, online providers may have more than one teacher teaching the same AP course. Yet, unlike traditional schools, online providers do not receive the annual AP Participation Survey in which schools report who their teacher or teachers are for each AP subject. Therefore, even if an online-only student indicated who was their online provider of instruction (and 35% of online-only students did not make this indication), because online schools do not submit the names of their teachers to the College Board, it is not possible to group specific online-only students with a specific online teacher. Therefore, the results of the study may be influenced by not being able to quantify an online classroom or teacher-effect (Bloom, 2005). This may result in smaller standard errors than those obtained had the teacher effect been estimated, thereby leading to a greater probability of rejecting the null hypothesis (Osborne, 2000).

While the analyses did not explicitly address the issue of the online classroom or teacher effect, an intraclass correlation was calculated for students who indicated their online provider (Raudenbush & Bryk, 2002). The intraclass correlation among students in an online class may be less than those in a traditional brick-and-mortar school due to the variation in their experiences. Traditional students attend at the same day and time in the same room, hear the same questions, and take in the same visual clues. Online students, although presumably responding to the same online text (via documents, emails, discussion boards, etc.), may not necessarily interact simultaneously with each other (if at all), may have more or less contact with their teacher, and generally have more independence in their studies. This investigation found the range of intraclass

correlations to be low (0.07 to 0.22) across online-only students' exams. Therefore, one can feel more comfortable that the variation in online student scores is more a factor of the student variation and less a factor of the variation due to an online provider of the instruction effect. Also one may have reasonable confidence in the standard errors from the regression analyses.

### ***Power and Sample Size***

In this study, it was anticipated that the overall impact of the group work and extended essay writing would be relatively small, given the presence of the independent variables and covariates. Yet, considering the sample size in each subject, it is probable that even relatively small effect sizes would be detectable.

Cohen (1988) created an  $f^2$  statistic that is a function of the proportion of the variance of the dependent variable that is accounted for by some source(s) (i.e., independent variables) over the proportion of error variance (pp. 407–411). In other words, the larger the variance accounted for by the independent variables, the larger the impact of the treatment (e.g., group work) on the outcome (e.g., student AP scores). Any  $f^2$  value between .02 and .14 is considered to have a small effect. Values of  $f^2$  greater than or equal to .15 are considered medium effects and values greater than .35 are considered large effects (Cohen, 1988, pp. 413–414). Given an alpha of .05, the total number of independent variables after dummy-coding, with power equal to or greater than .80, sample sizes for every exam studied should be sufficient to detect small effect sizes (minimum detectable effect sizes ranged from .02 to .04 across the nine exams).

## **Results**

### ***Online-Only Students***

Overall, from the 2003–2005 AP administrations only 12,379 of the 5.6 million exams were classified as taken by online-only students (see Table 6). Despite a seemingly low percentage of online-only students (0.2%), there is evidence that the move toward online courses is increasing. The growth in number of exams taken by online-only students from 2003 to 2005 (25%) outpaced the growth in exams taken by no-online students (21%).



**Table 6*****Number of AP Exams Administered 2003–2005 by Degree of AP Online Instruction Across All AP Exams***

	2003		2004		2005		2003-2005	
	# exams	%	# exams	%	# exams	%	# exams	%
All exams	1,698,179	100	1,840,837	100	2,054,179	100	5,593,195	100
No online	1,208,054	71	1,311,197	71	1,460,196	71	3,979,447	71
Online only	3,683	0.2	4,108	0.2	4,588	0.2	12,379	0.2

The online-only students in this study attended a wide variety of online schools. Students reported 28 specific online providers as the source of their online instruction. The most often mentioned were Apex Learning (2,471 exams), Florida Virtual School (970 exams), and Iowa Online Advanced Placement Academy (584 exams). In addition to having a variety of online providers, online-only students live in a variety of locations. Online students indicated a total of 2,440 unique attending institutions or high schools, from all 50 states and Washington, DC. The most frequently reported states were California (1,303 exams), Florida (1,160 exams), and Iowa (1,068 exams).

In general, online-only students received four to six hours of instruction each week (see Table B1 in Appendix B). Most students in the study were new to online AP courses. The median and mode across all subjects equaled 1 AP course taken online by the students in their academic lifetime. A few students have taken several online AP courses, some as many as eight (see Table B2). Students were more familiar with AP courses overall, however. While the mode for number of AP exams taken during their academic career in seven out of nine subjects equaled 1, the median across all subjects was at least 2, often 3, or greater (see Table B2). Some students have taken up to 17 AP exams.

The two online social sciences tests dealing with the United States (U.S. History and U.S. Government and Politics) both were comprised nearly equally of female and male students. Online female students outnumbered their male classmates in English Language and Composition, English Literature and Composition, and Psychology. Conversely, a greater percentage of online males were found in Microeconomics, Macroeconomics, Calculus AB, and Statistics (see Table B3).

The majority of online students in this study were White. The percentage reporting White as their race ranged from 67% (Macroeconomics and Calculus AB) to 81% (Psychology). The greatest percentage of traditionally underrepresented minority (non-White and non-Asian American or Pacific Islander) students were found in English Literature and Composition and U.S. Government and Politics (22% each), while Psychology and Statistics had the smallest percentage (12% each; see Table B3).

For most subjects the majority of online students from 2003–2005 were seniors in high school. However, U.S. History and English Language and Composition both had a majority of their students as juniors (67% and 62%, respectively). Additionally, U.S. History was the only subject to have over 10% of its students listed as sophomores (see Table B4).

Although the focus of this report is not on comparing the AP performance of online-only and no-online students, it is beneficial to get an understanding of the differences in demographics between the relatively new and unknown online population of AP students when compared to the historic and known traditional AP students. Because of the difficulty in matching individual records from one year to the next, a comparison was made of the 2005 online-only and no-online students only.

Of the nine subjects studied, three subjects had a statistically significant smaller proportion of female students in the online classroom (online-only students) when compared to the traditional classroom (no-online students). In 2005, about half of the online-only students in U.S. History (49%) were female; by comparison 56% of the no-online U.S. History students were female. For both math courses (Calculus AB and Statistics) fewer than 4 in 10 online-only students were female, while in the traditional classroom, roughly half of the no-online students were female. The remaining subjects did not differ with regard to the gender makeup (see Table B5).

The racial composition of online-only classrooms differed from traditional no-online classrooms. Most online-only courses had higher percentages of Black or African American students and fewer percentages of Asian American or Pacific Islander students and Mexican-American students than in traditional no-online courses. Only Microeconomics and Statistics reported no significant differences when viewing the percent of students by race across online-only and no-online classrooms. U.S. History stood out as the only subject where the percentage of online-only students was smaller for each non-White category compared to the percentage of no-online students (see Table B6).

In each exam significant differences were found across education level when comparing online-only students to no-online students. Generally, the online-only students in 2005 were younger than the no-online students. For six out of the nine exams, a greater percent of online-only students than no-online students reported to be juniors or sophomores. Conversely, in U.S. History, English Language and Composition, and Statistics a greater percentage of seniors were found online-only than in the no-online classrooms (see Table B7).

### ***Frequency and Impact of Group Work***

The frequency with which students participated in group work varied by subject. The majority of online students in six out of the nine subjects indicated that they never participated in group work in their AP class. In general, students in the math courses (Calculus AB and Statistics) and the Englishes (English Language and Composition and English Literature and Composition) most often worked in groups. The social sciences (Microeconomics, Macroeconomics, U.S. Government and Politics, and Psychology) had larger percentages of students replying they “never” worked in groups (see Table 7).

Each of the 27 regression models (9 exams \* 3 dependent variables) with the PSAT/NMSQT flag, but not the PSAT/NMSQT score, produced a statistically significant overall  $F$ -value. On average, models that controlled for group work, the PSAT/NMSQT flag only, and the remaining independent variables accounted for roughly one-quarter of the variance in the three dependent variables (free response score, multiple-choice score, and composite score). The median and mean total  $R^2$  equaled .23, with a standard deviation of .05. With the exception of U.S. History's free response scores, for each of the remaining 26 models a significant partial  $R^2$  for the set of group work, dummy variables were found.

Similarly, the overall  $F$ -value was significant for every regression that estimated the impact of group work on the three dependent variables after controlling for the PSAT/NMSQT score as well as the PSAT/NMSQT flag. On average, these models accounted for approximately half of the variance in the three dependent variables. The median and mean total  $R^2$  equaled .50, with a standard deviation of .10. For 12 of these 27 regression models a significant partial  $R^2$  for the set of group work, dummy variables were found.

### **Table 7**

#### ***Frequency of Group Work by AP Exam for Online-Only Students***

AP Exam	Never		A few times a year		Once/twice a month		Once/twice a week		About every day		Invalid response	
	<i>N</i>	Row %	<i>N</i>	Row %	<i>N</i>	Row %	<i>N</i>	Row %	<i>N</i>	Row %	<i>N</i>	Row %
U.S. History	543	50	172	16	123	11	132	12	88	8	31	3
Microeconomics	575	66	76	9	63	7	88	10	66	8	4	0
Macroeconomics	657	69	84	9	72	8	85	9	54	6	6	1
Eng. Lang. & Composition	383	45	192	23	127	15	90	11	47	6	11	1
Eng. Lit. & Composition	459	49	190	20	98	10	113	12	62	7	13	1
U.S. Government & Politics	866	67	150	12	85	7	104	8	70	5	10	1
Calculus AB	242	44	108	19	40	7	46	8	88	16	31	6
Psychology	945	74	82	6	60	5	108	9	65	5	9	1
Statistics	280	53	92	18	44	8	48	9	60	11	0	0

Only a few exams showed evidence of any impact from group work on the three AP scores. The impact in each of these few models was minimal. If the effect size ( $f^2$ ) value equals .02 it is, in Cohen’s words, “just barely escaping triviality” (Cohen, 1988, p. 413). The largest effect sizes found for the impact of group work on AP performance ranged from .02 to .05. These effect sizes were found on the English Literature and Composition, Calculus AB, and Statistics exams. All nine of the models with effect sizes greater than or equal to .02 were of the kind where the PSAT/NMSQT score was not included among the independent variables (PSAT Flag Only). When PSAT/NMSQT score was also controlled for, no model had an effect size greater than .01.

If there was to be an impact, minor though it was, it was found when students worked most frequently in groups. Four of the nine models with small effect sizes had students who worked “about every day” in groups perform lower than students who “never” worked in groups. Similarly, students without group work experience scored higher than those who worked in groups “once/twice a week” in four models.

The following sections will report findings on the three AP exams where effect sizes greater than or equal to .02 were found.

*English Literature and Composition.* Including only the PSAT/NMSQT flag along with the other independent variables and group work variables produced significant models for all three English Literature and Composition scores [free response:  $F(16, 905) = 5.18, p < .01$ ; multiple-choice:  $F(16, 905) = 11.73, p < .01$ ; composite:  $F(16, 905) = 10.66, p < .01$ ]. Group work dummy variables explained 2% of the variance in the free response score, 4% in the multiple-choice score, and 3% in the composite score after controlling for the variance explained by the other variables in the model. The total  $R^2$  for the three models ranged from .22 (free response) to .28 (multiple-choice and composite).

Nearly all of the group work dummy variables achieved statistical significance in the three models. The exception was students who replied they worked in groups “a few times a year” (see Table C1 to Table C3). However only three group work dummy variables reached the level of effect size that is considered small. In each case, the effect size equaled .02. The 175 students who worked in groups “once/twice a week” or “about every day” on average had lower multiple-choice scores than students who never worked in groups. Similarly, the 113 students who worked in groups weekly tended to score lower than students without group work experience on the composite score.

*Calculus AB.* Grouped group work dummy variables achieved statistical significance in all three score models after controlling for the variance associated with the other independent variables and only PSAT/NMSQT flag [free response:  $F(16, 507) = 7.63, p < .01$ ; multiple-choice:  $F(16, 507) = 10.27, p < .01$ ; composite:  $F(16, 507) = 9.53, p < .01$ ]. The partial  $R^2$  values were the highest found for any of the group work models that included the PSAT/NMSQT flag (free response .05, multiple-choice .06, composite .06). Total  $R^2$  values for the models ranged from .17 (free response) to .24 (multiple-choice).

Both dummy variables created for students who reported working in groups almost daily and “a few times a year” had scores significantly lower than other students who never worked in groups (see Table C4 to Table C6). Yet, it was only the dummy variables for the 88 students who worked almost daily in groups where effect sizes equal to or above .02 were found. On the free-response score the effect size equaled .04, while on both the multiple-choice and composite scores, the effect size reached .05, the largest found in the study.

*Statistics.* The amount of variance in the Statistics free response, multiple-choice, and composite scores explained by group work variables, after controlling for the other independent variables and PSAT/NMSQT flag only, all reached statistical significance. All three partial  $R^2$  values equaled .03 and had a significant F-values [free response  $F(16, 507) = 5.65, p < .01$ ), multiple-choice  $F(16, 507) = 5.01, p < .01$ , composite  $F(16, 507) = 5.93, p < .01$ ]. The total  $R^2$  values ranged from .25 (multiple-choice) to .28 (composite).

Dummy variables for working in groups “a few times a year,” “once/twice a month,” and “once/twice a week” were significant in all three outcome measures (see Table C7 to Table C9). Only the dummy variable for working in groups weekly reached a small effect size (.02), and then only on two scores. Statistics students who worked in groups a few times per week ( $n = 48$ ) had both lower free response and composite scores than students who never worked in groups.

***Frequency and Impact of Extended Essay Writing***

The frequency of students indicating they wrote extended essays as part of their online AP coursework varied by subject. The majority of online students in five out of the nine subjects indicated that they participated in extended essay writing in their AP class at least “a few times a year.” The math (Calculus AB, Statistics) or more quantitative (Microeconomics, Macroeconomics) courses had the majority of their students indicating that they never wrote extended essays, whereas the Englishes (English Language and Composition, English Literature and Composition) and less quantitative social sciences (U.S. History, U.S. Government and Politics, Psychology) were more likely to have students write extended essay in their respective courses “once/twice a week” or “once/twice a month” (see Table 8).

Across all nine exams, writing extended essays did not have a regular, nor appreciable impact on students’ AP performance. Like in the group work analysis, in the cases when there were significant partial  $R^2$  values for the extended essay writing dummy variables, the corresponding effect sizes were found to be minor. In fact, every model with a significant partial  $R^2$  value had an effect size of less than .02.

**Table 8**

***Frequency of Writing Extended Essays by AP Exam for Online-Only Students***

AP exam	Never	A few times a	Once/twice a month	Once/twice a week	About every	Invalid response
---------	-------	---------------	--------------------	-------------------	-------------	------------------

	year						day					
	<i>N</i>	Row %	<i>N</i>	Row %	<i>N</i>	Row %	<i>N</i>	Row %	<i>N</i>	Row %	<i>N</i>	Row %
U.S. History	48	4	150	14	322	30	531	49	26	2	12	1
Microeconomics	569	65	116	13	88	10	84	10	14	2	1	0
Macroeconomics	498	52	152	16	159	17	124	13	20	2	5	1
Eng. Lang & Composition	37	4	66	8	264	31	451	53	24	3	8	1
Eng. Lit. & Composition	47	5	98	10	278	30	468	50	38	4	6	1
U.S. Government & Politics	115	9	193	15	337	26	602	47	37	3	1	0
Calculus AB	480	86	21	4	14	3	20	4	16	3	4	1
Psychology	210	17	234	18	411	32	395	31	19	1	0	0
Statistics	430	82	47	9	19	4	16	3	10	2	2	0

## Discussion

### *Limitations of the Study*

This study identified students as online only according to their answers to two background questions. However, as is true in traditional classrooms, there exists a wide disparity across how online courses are structured, what level of interaction is expected, and what technologies are involved (e.g., video, Flash animation, etc.). Combining all online classrooms into one category gives the impression of uniformity when in fact there may be a great variety of online experiences.

There can also be a number of interpretations for the background questions of primary interest: group work and extended essay writing. The determination of frequency of group work came from the single background question asked of each examinee. In that question, group work is only defined as working in “pairs or small groups.” It does not elaborate on the intensity of the group work experience. One student may consider helping out a friend with their homework as

working in pairs, while another student may have considered the team-project as their group work. The extended essay writing question was a bit more specific in that it defined the term as “500 words or more.” Still, even with that definition, students may have differed in their understanding of extended essay writing.

Care therefore should be taken when reviewing these findings given that there are differences within online courses and the students’ interpretations of the group work and extended essay writing background questions. These differences, as is true in most research, may have a profound impact on results, especially when sample sizes are small. This study is also limited by the following factors:

- Given that this evaluation is not based upon data arising from a randomized experiment, it is susceptible to violations of internal and external validity. For example there are concerns regarding how students found their way to the online course (selection), changes in students with regard to the passage of time (maturation), and differential loss of students from the various group work/essay writing conditions (mortality; Campbell, & Stanley, 1966).
- Students self-reported whether or not their only source of instruction for the AP course was online, rather than records being gathered from the online-provider course rosters. It is possible that not all online students are accounted for.
- The students in this study are taking advanced coursework while still in high school. It is likely then that they differ from other online high school students who opt not to take AP courses. The findings from this study may not then apply to all online high school students.
- Academic performance in this study is measured by the students' achievement on the AP exams, which are given at the end of the academic year. However, some students who have gone through an AP course may choose not to take the AP exam. Other students who began the course may have dropped out after a period of time. Findings therefore should not be projected on all online students who take AP courses.

### ***Online AP Students***



Overall, this study provided new information regarding the growth of AP online, where it is being offered and who is taking it. While the vast majority of the students taking AP are traditional students, it is important to note that the number of exams taken by online students is growing at a faster rate than the number of exams taken by traditional students. From 2003 to 2005, the total number of AP exams taken grew by 21%. Even more noticeable is that over the same time period, the number of AP exams taken by online-only students grew by 25%. Online high schools may therefore become an increasingly larger subset of schools that offer AP.

Online high schools do not exist in an isolated region or part of the nation. The greatest percentage of AP exams taken by online-only students came from the West Coast, the southeastern corner of the nation, and the Midwest. However, every state in the Union was represented in the study.

One in seven high schools relied on online institutions to supplement their AP course offerings. Between the years 2003 and 2005, almost 17,000 high schools administered AP exams. Of those 17,000 high schools, online-only students indicated attending over 2,400 of them (14%).

Online high schools appear to be partly successful in providing access to AP courses for non-White students. One of the predominant reasons for the establishment of online high schools was to provide courses to students who historically have not had access to a wide variety of curriculum. Often these students are non-White and living in urban or rural parts of states. In fact, most of the online courses studied had greater percentages of Black or African American students than in the traditional classrooms. However, the lack of Hispanic students in many of these courses raises concern that these students may not have access to online courses or may not be taking advantage of them where access exists.

### ***How Frequently Are Online-Only AP Students Working in Pairs or Small Groups and Does It Impact Their AP Exam Performance?***

Online high schools do not appear to differ from traditional high schools in the importance they place on having their students participate in group work. Study results show online-only students rarely indicated working in groups. These results are similar to findings from the NAEP data and other studies that revealed that group work is not pervasive in traditional high school classrooms (Langer, 2001, see also Tables A1–A7).

This lack of time spent working in groups does not appear to have an impact on the online-only students' subsequent performance on the AP exam. In most subjects studied, working in groups

had no impact on the students' AP scores. For the very few cases where an impact was found, working in groups had a minimal but negative effect on student performance, meaning that students who worked in groups had lower scores than those who did not. However, these results appeared on the more extreme end of the group work continuum. It was only for the few who worked in groups very often, either weekly or daily, where scores were slightly lower on three exams.

There may be a few plausible explanations for this occurrence. First, on the three exams where this effect was found (English Literature and Composition, Calculus AB, and Statistics) a student having a deep understanding of specific content is very important. If the students work so regularly in groups that there is little to no individual work, those students may not have internalized the knowledge because they had never had to as part of their class work. Furthermore, all three of these subjects are highly analytical in nature. Although previous studies have reported gains in analytical thinking skills through the use of group work, excessive group work may not provide the opportunity for the students to refine these skills independently. It is possible to consider that working in groups may expose students to various ways of approaching problems that lead to gains in analytical skills. However, if the student is not given the opportunity to independently cultivate these skills, that exposure may not take root.

Another explanation for slightly lower performance on these three exams for students working in groups is the possibility that a group can succeed while the individuals do not. When pooling knowledge, talent, and abilities the students working in groups may be able to successfully fill in the holes or gaps for each other to complete the project. However, if the students learn that their group members will be able to compensate for their shortcomings and choose not to develop these areas within themselves, when it comes time to take an examination as individuals, such as the AP exam, the students' weaknesses or gaps in knowledge will be more easily identified and reflected in their scores. Fisher, Thompson, and Silverberg (2005) noted that the success of online group work "is directly related to active participation and the intrinsic or extrinsic goals and needs of the learner" (p. 217). Even students who are motivated enough to take on the challenge of online AP coursework while a junior (or even a sophomore) may not necessarily see the implication of letting the others in the group handle it. It is for this reason that many researchers when looking at group work have encouraged instructors to be sure to make each student accountable, to be able to identify each student's contribution to the group, and to supply incentives for each student's participation (Meyers, 1997; Vermette, 1995).

While the AP course is designed to expose high school students to the rigorous coursework and assignments they would face on a college campus, the exams have the “purpose of assessing the students’ preparation in the subject” (College Board, 2007b). As such, the exams may not necessarily measure explicitly interpersonal skills, sense of responsibility, and confidence—all skills reported to have been improved by students working in groups. Boud, Cohen, and Sampson (1999) stated that to understand the benefits of peer learning or group work, it is peer learning that must be assessed. Another frequently reported benefit to group learning, that being the decrease in students’ sense of isolation, is also not measured by the AP exam.

Although there is no evidence in this study to suggest group work improves AP scores, it may help online students, who are by the nature of online classrooms physically removed from their classmates, feel more connected to one another. With a larger percentage of students online being younger and the fact that for most this is their first online experience, activities that bring the students together may have a positive influence, but not on AP scores.

### ***How Frequently Are Online-Only AP Students Writing Extended Essays and Does It Impact Their AP Exam Performance?***

Extended essay writing is a frequent activity in most online AP English courses. Students in the less quantitative social sciences wrote extended essays monthly or weekly, while students in the math and more quantitative social sciences wrote the least, if at all. These findings are very similar to what was seen when viewing NAEP data for all U.S. high school seniors (see Tables A8-A16).

When controlling for demographics, AP experience, and PSAT/NMSQT data, no appreciable impacts on achievement were found for writing extended essays. One possible explanation is the disparity between what constitutes extended essay writing in high schools and what it means in colleges. Extended essay writing was defined in the AP student questionnaire as a minimum of 500 words (approximately two pages). Yet a recent study conducted by *The Chronicle for Higher Education* revealed that college faculty require much longer essays (Sanoff, 2006). Greater than 70% of the nearly 1,100 college faculty members who replied to the survey indicated they require students to write papers of at least five pages at least occasionally during the course of the semester. Perhaps, the writing assignments given to the AP online students were not of great enough length for them to affect achievement. Furthermore, the quality of the writing by

the students and the degree of difficulty of the tasks given by the high school teachers may not be up to college level. *The Chronicle* survey reported 54% of the faculty said students are not well prepared for college-level writing. Contrasting that figure, only 10% of the 746 high school teachers who replied to the survey agreed that students are not well prepared for college-level writing.

## **Implications**

### ***For Online-Instruction Providers***

One of the goals of this research was to provide useful feedback to online providers of instruction as to what works in their AP courses. Findings from this study can help inform their course designs in a few ways.

First, it is important to understand that more than in traditional AP courses, online-only students tend to be younger. They also are new to the online classroom experience. Since the AP is designed to simulate college-level academic experience while still in high school, extra time may need to be given to assimilating these students into the workings of online courses in general and the high level of expectations for AP courses in particular.

When developing group work activities care may need to be taken regarding the frequency with which group work activities are assigned, if the primary goal is to achieve higher AP exam scores. Again, the results from this study were highly consistent in that group work did not appear to have an impact on AP scores. However, in a few cases, particularly in subjects where there is a great need for students to internalize specific content, such as formulas and equations, excessive group work may not be helpful in the online forum. When students have to take assessments where the answers are more clearly right or wrong, disproportionate use of group work may limit the opportunities students have to work on their own and internalize the content. While this study did not examine how the students worked within their groups, it does seem reasonable that making sure each student is accountable to the group and that their individual contributions can be measured are appropriate ways to help ensure that the individuals are learning within the group setting.

Although no impact was found on AP scores when looking at the frequency with which online students wrote extended essays, online instructors should not abandon the practice, particularly in the face of recent studies that report college faculty wanting to see more writing in

high school, not less. The question asked of AP students identified writing assignments of at least 500 words. If the colleges require students to write closer to 1,500 words, that may be the more appropriate length, especially for AP students.

### ***For Future Research***

Considering the newness of online high schools, and the lack of research into their performance (as defined any number of ways), it is clear that further research is needed. A general topic for further investigation is describing the variety of online course structures and their prevalence. Also, studies would be beneficial that examine why more African American students are taking online courses, and why fewer Asian and Hispanic students are venturing online.

Future studies into group work and extended essay writing, regardless of whether or not the students are online only, would benefit from more thoroughly focused questions on these issues. In this study there was only one group work question and one extended essay writing question available, designed and asked by the College Board, on their AP exams. The reliability of any one single item is generally very low, plus, the questions used here were open to considerable variation in interpretation on the parts of the students. Therefore, to be more helpful several new items would be needed that are more specific in their definitions of group work and extended essay writing.

The impact of group work on student isolation, a frequently cited benefit, should also be investigated, particularly given that these students are younger than their traditional AP peers and are going online for the first time to take these courses. As well, it would be beneficial to research how various ways of holding students accountable for their online group work impacts their performance and development on the skills mentioned previously.

Investigation into the forms and frequency of feedback from online AP instructors to students could help add to the understanding of how online students could learn to write better. It is reasonable to assume that engaging in writing without feedback would provide little benefit to the students. Consequently, how online teachers reply to their students' work may play a part in their students' improvement.

### **References**

Adelman, C. (2006). *The toolbox revisited: Paths to degree completion from high school through college*. Washington, DC: U.S. Department of Education.

- Angelo, T. A. (1995). Classroom assessment for critical thinking. *Teaching of Psychology*, 22(1), 6–7.
- Baskin, C. (2001, December). *Using Kirkpatrick's four-level-evaluation Model to explore the effectiveness of collaborative online group work*. Paper presented at the annual conference of the Australasian Society for Computers in Learning in Tertiary Education. Retrieved March 3, 2007, from <http://www.ascilite.org.au/conferences/melbourne01/pdf/papers/baskinc.pdf>
- Bauer, J. F., & Anderson, R. S. (2000). Evaluating students' written performance in the online classroom. *New Directions for Teaching and Learning*, 84, 65–71. Retrieved from <http://proxy.library.upenn.edu:8117/cgi-bin/fulltext/101523976/PDFSTART>
- Bigbie, C. L., & McCarroll, W. J. (2000). *The Florida high school evaluation 1999-2000 report*. Tallahassee, FL: Florida State University College of Education Center for the Study of Teaching and Learning.
- Blomeyer, R. (2002). Virtual schools and e-learning in K-12 environments: Emerging policy and practice. *NCREL Policy Issues*, 11, 1–4, 7–10.
- Bloom, H. S. (Ed.). (2005). *Learning more from social experiments*. New York: Russell Sage Foundation.
- Bottoms, G., & Bearman, A. (2000). *Improving reading and writing skills in language arts courses and across the curriculum* (High Schools That Work Research Brief). Atlanta, GA: Southern Regional Education Board.
- Boud, D., Cohen, R., & Sampson, J. (1999). Peer learning and assessment. *Assessment and Evaluation in Higher Education*, 24(4), 413–426.
- Breland, H. M., Bonner, M. W., & Kubota, M. Y. (1995). *Factors in performance on brief, impromptu essay examinations* (College Board Research Rep. No. 95-4). New York: The College Board.
- Camara, W. J., & Millsap, R. (1998). *Using the PSAT/NMSQT and course grades in predicting success in the Advanced Placement Program<sup>®</sup>* (College Board Research Rep. No. 98-4). New York: The College Board.
- Campbell, D. T., & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.

- Cavanagh, S. (2006, February 15). Higher proportion of seniors mastering AP. *Education Week*. Retrieved from <http://www.edweek.org/ew/articles/2006/02/15/23ap.h25.html?querystring=Mastering%20AP&levelId=2200>
- Chen, F., Elbaum, B., & Walsh, S. (1999). *Students speak: An analysis of student feedback on the virtual high school experience*. Retrieved March 3, 2007, from [http://www.govhs.org/vhsweb/press.nsf/a59a5b767cfb7e008525664f005b1a2c/7a4ea8090dde28b506256b66006fe453/\\$FILE/aerapaper.pdf](http://www.govhs.org/vhsweb/press.nsf/a59a5b767cfb7e008525664f005b1a2c/7a4ea8090dde28b506256b66006fe453/$FILE/aerapaper.pdf)
- Clark, T. (2000). *Virtual high schools: State of the states*. Macomb, IL: Western Illinois University, Center for the Application of Information Technologies.
- Clark, T. (2001). *A study of virtual schools in the United States*. Macomb, IL: Western Illinois University, Center for the Application of Information Technologies.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- College Board. (2007a). *AP research and data*. Retrieved from the College Board AP Central Web site: <http://apcentral.collegeboard.com/apc/members/program/research/index.html>
- College Board. (2007b). *Course and exam development*. Retrieved from the College Board AP Central Web site: [http://apcentral.collegeboard.com/apc/public/exam/about\\_exams/2024.html](http://apcentral.collegeboard.com/apc/public/exam/about_exams/2024.html)
- College Board. (2007c). *Exam data 2005*. Retrieved March 3, 2007, from the College Board AP Central Web site: [http://apcentral.collegeboard.com/apc/members/exam/exam\\_questions/47040.html](http://apcentral.collegeboard.com/apc/members/exam/exam_questions/47040.html)
- College Board. (2007d). *Exam scoring*. Retrieved March 3, 2007, from the College Board AP Central Web site: [http://apcentral.collegeboard.com/apc/public/exam/about\\_exams/1994.html#name3](http://apcentral.collegeboard.com/apc/public/exam/about_exams/1994.html#name3)
- Connor-Greene, P. A., & Murdoch, J. W. (2000). Does writing matter? Assessing the impact of daily essay quizzes in enhancing student learning. *Language and Learning Across the Disciplines*, 4(1), 16–21.
- Cook, G. (2002). The cyber charter challenge. *American School Board Journal*, 189(9). Retrieved March 3, 2007, from <http://www.asbj.com/2002/09/0902ASBJS2.pdf>

- Cooper, S. S. (2004). *Metacognition in the adult learner*. Retrieved March 3, 2007, from <http://www.lifecircles-inc.com/metacognition.htm>
- Dillon, J. (2005, September 29). Online program offers AP classes. *The Casper Star-Tribune*. Retrieved March 3, 2007, from <http://www.casperstartribune.net/articles/2005/09/29/news/wyoming/4f36d72ed898b1838725708b000179c1.txt>
- Doherty, K. (2002, May 9). Students speak out. *Education Week*. Retrieved from <http://www.edweek.org/sreports/tc02/article.cfm?slug=35florida.h2>
- Education Week. (2005a, August 10). *Evaluating online teachers is largely a virtual task*. Retrieved from <http://www.edweek.org/ew/articles/2005/08/10/44onlineeval.h24.html>
- Education Week. (2005b, September 8). *Ins and outs of virtual schools. Transcript from EdWeek live chat*. Retrieved March 3, 2007, from [http://www.edweek.org/chat/transcript\\_09\\_08\\_2005.html](http://www.edweek.org/chat/transcript_09_08_2005.html)
- Fisher, M., Thompson, G. S., & Silverberg, D. A. (2005). Effective group dynamics in e-learning: Case study. *Journal of Educational Technology Systems*, 33(3), 205–223.
- Interactive Educational Systems Design, Inc. (2002). *Final survey report: Online courses and other types of online learning for high school students*. Prepared for Apex Learning and Blackboard Inc.
- Jan, T. (2005, July 18). State high schools log on for AP courses. *The Boston Globe*. Retrieved from [http://www.boston.com/news/education/k\\_12/articles/2005/07/18/state\\_high\\_schools\\_log\\_on\\_for\\_ap\\_courses/?page=full](http://www.boston.com/news/education/k_12/articles/2005/07/18/state_high_schools_log_on_for_ap_courses/?page=full)
- Kear, K. (2004). Peer learning using asynchronous discussion systems in distance education. *Open Learning*, 19(2), 151–165.
- Langer, J. A. (2001). Beating the odds: Teaching middle and high school students to read and write well. *American Educational Research Journal*, 38(4), 837–880.
- Little, R. J. A., & Rubin, D. B. (1987). *Statistical analysis with missing data*. New York: John Wiley & Sons, Inc.
- Lockee, B., Moore, M., & Burton, J. (2001). Old concerns with new distance education research. *Educause Quarterly*, 2, 60–62.
- Lopez, W. J. (2003). Content delivery for a virtual high school. *T.H.E. Journal*, 31(1), 30–33. Retrieved March 3, 2007, from <http://www.thejournal.com/magazine/vault/A4470.cfm>
- Luria, A. R., & Yudovich, F. I. (1971). *Speech and the development of mental processes in the child*. Baltimore, MD: Penguin.



- Marcel, K. W. (2003). *Online Advanced Placement courses: Experiences of rural and low-income high school students*. Retrieved May 8, 2007, from the Western Consortium for Accelerated Learning Opportunities Web site:  
[http://www.wiche.edu/Policy/WCALO/documents/MarcelOnline\\_000.pdf](http://www.wiche.edu/Policy/WCALO/documents/MarcelOnline_000.pdf)
- Marzano, R. J. (1993). How classroom teachers approach the teaching of thinking. *Theory Into Practice*, 32(3), 154–160.
- Meyers, S. A. (1997). Increasing student participation and productivity in small-group activities for psychology classes. *Teaching of Psychology*, 24, 105–115.
- National Association of State Boards of Education. (2001). *Any time, any place, any path, any place: Taking the lead on e-learning policy*. Retrieved March 3, 2007, from  
[http://www.nasbe.org/Organization\\_Information/e\\_learning.pdf](http://www.nasbe.org/Organization_Information/e_learning.pdf)
- National School Boards Foundation. (2002). *Are we there yet?* Retrieved March 3, 2007, from  
<http://www.nsbfb.org/theyet/fulltext.htm>
- Osborne, J. W. (2000). Advantages of hierarchical linear modeling. *Practical Assessment, Research & Evaluation*, 7(1). Retrieved March 3, 2007, from  
<http://PAREonline.net/getvn.asp?v=7&n=1>
- Parker, A. (1996, October). *A distance education how-to manual: Recommendations from the field*. Paper presented at the WebNet 96 conference, San Francisco, CA.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage Publications.
- Russell, G. (2001, September). Virtual schools real concerns. *Electronic school*. Retrieved March 3, 2007, from <http://www.electronic-school.com/2001/09/0901virtualschools.html>
- Sanoff, A. P. (2006, March 10). What professors and teachers think: A perception gap over students' preparation. *The Chronicle for Higher Education*. Retrieved March 3, 2007, from  
<http://chronicle.com/weekly/v52/i27/27b00901.htm>
- Saunders, P., & Scialfa, C. T. (2003). The effects of pre-exam instruction on students' performance on an effective writing exam. *Written Communication*, 20(2), 195–212.
- Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, 19, 460–475.
- Scott, K. (2002). Writing improvement for all. *Journal of School Improvement*, 3(1). Retrieved March 3, 2007, from <http://www.ncacasi.org/jsi/2002v3i1/writing>

- University of California College Preparatory Initiative. (2002). *March status report*. Retrieved March 3, 2007, from [http://www.uccp.org/docs/March\\_Status.pdf](http://www.uccp.org/docs/March_Status.pdf)
- Vail, K. (2001, September). Online learning grows up. *Electronic School*. Retrieved March 3, 2007, from <http://www.electronic-school.com/2001/09/0901f1.html>
- Vail, K. (2002). A new kind of school. *American School Board Journal*, 189(9). Retrieved March 3, 2007, from <http://www.asbj.com/2002/09/0902ASBJS1.pdf>
- Vermette, P. J. (1995). Cooperative learning teams: Eight suggestions for the novice user. *The Clearing House*, 68(5), 278–281.
- Vrasidas, C., Zembylas, M., & Chamberlain, C. R. (2003). Complexities in the evaluation of distance education and virtual schooling. *Educational Media International*, 40(3), 201–208.
- Wood, C. (2005, April/May). Highschool.com. *Edutopia*, 1(4), 32–37.

## Notes

<sup>1</sup> The AP Program provides an opportunity for high school students to take college-level courses while in high school. At the end of the school year, students may opt to take the AP exam for that subject. Should students score high enough, they may be eligible to receive college-credit from the institution of their choice. To date, there are 34 AP exams in 22 subject areas offered by more than 15,000 schools around the world. In 2005, 2.1 million exams were taken by over 1.2 million students (College Board, 2007c).

<sup>2</sup> “When the free-response section of an exam contains two or more parts, those parts are weighted according to a value assigned to them by the Development Committee. This allows the committee to place more importance on certain skills to correspond to their emphasis in the corresponding college course curriculum. Weighting also comes into play when looking at the multiple-choice section in comparison to the free-response section. For each AP exam, there is a formula for combining the scores for the multiple-choice and free-response sections or subsections into a maximum weighted score (composite score). Once the weights have been decided and the free-response section scored, computing each student's composite score is a purely mechanical process and is done by computer” (College Board, 2007d).

## Appendix A

### NAEP Tables on Group Work and Extended Essay Writing

**Table A1**

*When You Have Reading Assignments in School, How Often Does Your Teacher Do Each of the Following? Ask You To Do a Group Activity or Project About What You Have Read. ]*

Year	Almost every day		1–2 times a week		1–2 times a month		Never or hardly ever	
	Avg. scale score	Row pct	Avg. scale score	Row %	Avg. scale score	Row %	Avg. scale score	Row %
1998	292 (2.8)	6% (0.3)	295 (1)	24% (0.7)	294 (0.7)	43% (0.6)	283 (1.2)	27% (0.7)
1994	281 (2.1)	7% (0.4)	286 (1)	24% (0.7)	291 (0.9)	42% (0.6)	287 (1.2)	27% (0.8)
1992	288 (1.8)	6% (0.4)	289 (1)	22% (0.5)	295 (0.8)	42% (0.6)	292 (0.9)	29% (0.7)

*Note.* Standard errors appear in parentheses. The NAEP Reading scale ranges from 0 to 500. Observed differences are not necessarily statistically significant. From *U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1998, 1994 and 1992 Reading Assessments, Grade 12.* [Data file]. Available from the NAEP Data Explorer Web site, <http://nces.ed.gov/nationsreportcard/nde/criteria.asp>

**Table A2*****For Your English Class This Year, How Often Do You Work in Pairs or Small Groups To Talk About Something That You Have Read?***

	At least once a week		Once or twice/month		Few times a year		Never or hardly ever	
	Avg. scale score	Row %	Avg. scale score	Row %	Avg. scale score	Row %	Avg. scale score	Row %
Total	289 (1.0)	34% (1.0)	289 (0.8)	36% (0.6)	288 (1.3)	14% (0.4)	279 (1.4)	15% (0.7)

*Note.* Standard errors appear in parentheses. The NAEP Reading scale ranges from 0 to 500. Observed differences are not necessarily statistically significant. From U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2002 Reading Assessments, Grade 12.

38

**Table A3*****How Often Does Your Teacher Ask You To Do Each of the Following? Work in Pairs or Small Groups To Discuss Your Writing***

	Almost every day		1–2 times a week		1–2 times a month		Never or hardly ever	
	Avg. scale score	Row %	Avg. scale score	Row %	Avg. scale score	Row %	Avg. scale score	Row %
Total	150 (1.5)	13% (0.5)	152 (0.8)	30% (0.8)	153 (0.7)	31% (0.6)	148 (0.9)	26% (0.9)

*Note.* Standard errors appear in parentheses. The NAEP Writing scale ranges from 0 to 300. Observed differences are not necessarily statistically significant. From U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1998 Writing Assessments, Grade 12.

**Table A4**

*When You Write a Paper or Report for School This Year, How Often Do You Do Each of the Following? Work With Other Students in Pairs or Small Groups To Discuss and Improve Your Paper*

	Almost always		Sometimes		Never or hardly ever	
	Avg. scale score	Row %	Avg. scale score	Row %	Avg. scale score	Row %
Total	150 (1.2)	23% (0.5)	150 (0.9)	50% (0.5)	148 (1.0)	27% (0.5)

*Note.* Standard errors appear in parentheses. The NAEP Writing scale ranges from 0 to 300. Observed differences are not necessarily statistically significant. From U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2002 Writing Assessments, Grade 12.

39

**Table A5**

*When You Do Mathematics in School, How Often Do You Solve Mathematics Problems With a Partner or in Small Groups?*

Year	Almost every day		1–2 times a week		1–2 times a month		Never or hardly ever	
	Avg. scale score	Row %	Avg. scale score	Row %	Avg. scale score	Row %	Avg. scale score	Row %
2000	309 (1.4)	21% (0.8)	303 (1.3)	32% (0.7)	302 (1.3)	18% (0.6)	291 (1.1)	29% (0.9)
1996	304 (1.6)	20% (0.8)	304 (1.3)	30% (0.8)	305 (1.3)	21% (0.7)	296 (1.2)	30% (0.9)

*Note.* The NAEP Mathematics scale ranges from 0 to 500. Observed differences are not necessarily statistically significant. From U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2000, 1996 Mathematics Assessments, Grade 12.

**Table A6*****When You Study Science in School, How Often Do You Work With Other Students on a Science Activity or Project?***

	Almost every day		1–2 times a week		1–2 times a month		Never or hardly ever	
	Avg. scale score	Row %	Avg. scale score	Row %	Avg. scale score	Row %	Avg. scale score	Row %
2000	153 (1.5)	16% (0.8)	155 (1)	30% (0.6)	150 (0.9)	27% (0.5)	132 (1.1)	27% (0.9)
1996	160 (1.2)	14% (0.6)	162 (1.1)	30% (0.8)	153 (1.1)	25% (0.6)	136 (0.9)	31% (0.8)

*Note.* Standard errors appear in parentheses. The NAEP Science scale ranges from 0 to 300. Observed differences are not necessarily statistically significant. From U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2000, 1996 Science Assessments, Grade 12.

40

**Table A7*****When You Work on Science Experiments or Investigations, Do You Usually Work With Other Students?***

	Yes		No		No experiments		Not taking science	
	Avg. scale score	Row %	Avg. scale score	Row %	Avg. scale score	Row %	Avg. scale score	Row %
Total	156 (0.9)	52% (1.0)	135 (2.6)	4% (0.2)	137 (3.1)	3% (0.2)	138 (1.0)	41% (1.0)

*Note.* Standard errors appear in parentheses. The NAEP Science scale ranges from 0 to 300. Observed differences are not necessarily statistically significant. From U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2000 Science Assessments, Grade 12.

**Table A8*****In Mathematics Class, How Often Do You Do Each of the Following? Write Reports or Do Mathematics Projects***

	Almost every day		Several times a week		About once a week		Less than once/week		Never	
	Avg. scale score	Row %	Avg. scale score	Row %	Avg. scale score	Row %	Avg. scale score	Row %	Avg. scale score	Row %
Total	----	1% (0.2)	275 (4.3)	2% (0.4)	287 (4.5)	6% (0.6)	295 (1.9)	20% (1.1)	296 (1.2)	71% (1.2)

*Note.* Standard errors appear in parentheses. The NAEP Mathematics scale ranges from 0 to 500. Observed differences are not necessarily statistically significant. From U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1990 Mathematics Assessments, Grade 12.

41

**Table A9*****Do You Do Any of the Following When You Study Social Studies? Write Reports***

	Yes		No		I don't know	
	Avg. scale score	Row %	Avg. scale score	Row %	Avg. scale score	Row %
Total	152 (0.8)	67% (0.9)	150 (1.1)	29% (0.9)	128 (4.1)	3% (0.3)

*Note.* Standard errors appear in parentheses. The NAEP Civics scale ranges from 0 to 300. Observed differences are not necessarily statistically significant. From U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1998 Civics Assessments, Grade 12.



**Table A10*****When You Study Science in School, How Often Do You Prepare a Written Science Report?***

Year	Almost every day		1–2 times a week		1–2 times a month		Never or hardly ever	
	Avg. scale score	Row %	Avg. scale score	Row %	Avg. scale score	Row %	Avg. scale score	Row %
2000	136 (3.3)	2% (0.2)	152 (1.7)	13% (0.7)	151 (0.9)	37% (0.8)	143 (1.0)	48% (1.2)
1996	135 (3.3)	2% (0.2)	158 (1.8)	11% (0.6)	156 (0.9)	34% (0.7)	147 (0.8)	53% (0.9)

*Note.* Standard errors appear in parentheses. The NAEP Science scale ranges from 0 to 300. Observed differences are not necessarily statistically significant. From U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2000, 1996 Science Assessments, Grade 12.

42

**Table A11*****About How Often Do You Do Each of the Following in Your Class? Write Up the Results of the Science Experiment or Investigation You Designed***

Year	1–2 times a month		Less than once/month		Never		Not taking science	
	Avg. scale score	Row %	Avg. scale score	Row %	Avg. scale score	Row %	Avg. scale score	Row %
2000	148 (1.9)	10% (0.5)	154 (1.5)	17% (0.5)	148 (0.9)	45% (0.9)	142 (1.0)	27% (0.8)

*Note.* Standard errors appear in parentheses. The NAEP Science scale ranges from 0 to 300. Observed differences are not necessarily statistically significant. From U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2000 Science Assessments, Grade 12.

**Table A12**

***How Often Do You Write Something That Is at Least a Paragraph in Length for Each Class?***

Class	At least once a week		Once or twice/month		A few times a year		Never or hardly ever	
	Avg. scale score	Row %	Avg. scale score	Row %	Avg. scale score	Row %	Avg. scale score	Row %
English	154 (0.8)	74% (0.6)	144 (1.0)	18% (0.5)	120 (1.7)	5% (0.2)	112 (1.8)	4% (0.3)
Math	138 (1.7)	8% (0.3)	147 (1.6)	11% (0.3)	157 (1.2)	16% (0.4)	149 (0.8)	66% (0.7)
Science	154 (1.1)	20% (0.6)	154 (1.0)	27% (0.5)	150 (1.1)	17% (0.4)	143 (0.9)	36% (0.8)
Soc. studies/history	156 (0.9)	40% (0.8)	152 (1.0)	29% (0.5)	142 (1.4)	13% (0.4)	136 (1.2)	17% (0.6)

*Note.* Standard errors appear in parentheses. The NAEP Writing scale ranges from 0 to 300. Observed differences are not necessarily statistically significant. From U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2002 Writing Assessments, Grade 12.

**Table A13***This Year in School, How Often Have You Been Asked to Write Long Answers to Questions on Tests or Assignments?*

Assessment	Year	At least once a week		1–2 times a month		1–2 times a year		Never	
		Avg. scale score	Row %	Avg. scale score	Row %	Avg. scale score	Row %	Avg. scale score	Row %
Geography	1994	282 (1.4)	16% (0.5)	280 (1.3)	15% (0.6)	281 (1.5)	15% (0.6)	288 (0.8)	54% (0.9)
Reading	2002	293 (1.0)	31% (0.6)	291 (0.8)	41% (0.6)	281 (1.3)	22% (0.6)	264 (2.2)	7% (0.3)
Reading	1998	297 (1.0)	44% (0.8)	292 (1.0)	39% (0.6)	281 (1.2)	13% (0.5)	257 (2.6)	4% (0.2)
Reading	1994	294 (1.1)	45% (0.9)	289 (0.9)	38% (0.7)	274 (1.6)	12% (0.5)	255 (2.4)	4% (0.2)
Reading	1992	299 (0.8)	46% (0.9)	292 (0.8)	37% (0.7)	281 (1.4)	13% (0.5)	267 (2.6)	5% (0.4)
Science	2000	152 (1.3)	27% (0.8)	154 (1.1)	27% (0.6)	150 (1.3)	14% (0.5)	140 (1.1)	33% (0.8)
Science	1996	156 (0.9)	28% (0.8)	158 (1.0)	26% (0.5)	150 (1.4)	14% (0.4)	143 (1.2)	32% (0.7)
U.S. History	2000	290 (0.9)	26% (0.8)	289 (1.2)	28% (0.8)	284 (1.4)	13% (0.6)	284 (1.2)	32% (1.3)

*Note.* Standard errors appear in parentheses. The NAEP scale ranges by assessment: Geography from 0 to 500, Reading from 0 to 500, Science from 0 to 300, and U.S. History from 0 to 500. Observed differences are not necessarily statistically significant. From U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 Geography Assessments; 2002, 1998, 1994 and 1992 Reading Assessments; 1996 Science Assessments; 1994 U.S. History Assessments, Grade 12.

**Table A14**

*For Your English Class So Far This Year, How Many Times Have You Written a Report or Paper About Something That You Have Read (For Example, a Book Report)?*

	6 times or more		4 or 5 times		2 or 3 times		Once		Never	
	Avg. scale score	Row %	Avg. scale score	Row %	Avg. scale score	Row %	Avg. scale score	Row %	Avg. scale score	Row %
Total	295 (1.1)	25% (0.7)	292 (1.1)	21% (0.5)	286 (0.9)	32% (0.7)	279 (1.3)	13% (0.4)	277 (1.9)	9% (0.4)

*Note.* Standard errors appear in parentheses. The NAEP Reading scale ranges from 0 to 500. Observed differences are not necessarily statistically significant. From U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2002 Reading Assessments, Grade 12.

**Table A15**

*How Often Are Papers of the Following Lengths Assigned in Your English Class?*

Pages	Almost every day		1–2 times a week		1–2 times a month		Never or hardly ever	
	Avg. scale score	Row %	Avg. scale score	Row %	Avg. scale score	Row %	Avg. scale score	Row %
One to two	139 (1.6)	7% (0.3)	154 (0.9)	33% (0.6)	154 (0.7)	47% (0.6)	137 (1.2)	13% (0.5)
Three or more	126 (2.0)	3% (0.1)	144 (1.5)	9% (0.3)	157 (0.7)	49% (0.7)	146 (0.8)	39% (0.7)

*Note.* Standard errors appear in parentheses. The NAEP Writing scale ranges from 0 to 300. Observed differences are not necessarily statistically significant. From U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1998 Writing Assessments, Grade 12

**Table A16***When You Do Mathematics in School, How Often Do You Do Each of the Following? Write Reports or Do Mathematics Projects*

	Almost every day		1–2 times a week		1–2 times a month		Never	
	Avg. scale score	Row %	Avg. scale score	Row %	Avg. scale score	Row %	Avg. scale score	Row %
1996	279 (4.6)	2% (0.2)	287 (2.9)	4% (0.3)	301 (1.4)	23% (1.0)	304 (1.1)	71% (1.1)
1992	269 (3.4)	1% (0.1)	279 (3.7)	3% (0.2)	297 (1.9)	14% (0.6)	301 (0.9)	82% (0.7)

*Note.* Standard errors appear in parentheses. The NAEP Mathematics scale ranges from 0 to 500. Observed differences are not necessarily statistically significant. From U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992 and 1996 Mathematics Assessments, Grade 12.

**Appendix B**  
**AP Tables for Online Instruction**

**Table B1**

*Hours of Online Instruction Received per Week by AP Exam for Online-Only Students 2003–2005*

AP exam	<1 hr/wk		1–3 hrs/wk		4–6 hrs/wk		>6 hrs/wk	
	<i>N</i>	Row %	<i>N</i>	Row %	<i>N</i>	Row %	<i>N</i>	Row %
U.S. History	129	12	236	22	456	42	268	25
Microeconomics	55	6	222	25	405	46	190	22
Macroeconomics	70	7	292	30	408	43	188	20
English Language & Composition	106	12	189	22	379	45	176	21
English Literature & Composition	95	10	241	26	395	42	204	22
U.S. Government & Politics	110	9	327	25	557	43	291	23
Calculus AB	71	13	129	23	248	45	107	19
Psychology	154	12	306	24	540	43	269	21
Statistics	35	7	149	28	232	44	108	21

**Table B2***Average Number AP Exams Lifetime and Number AP Courses Taken Online by AP Exam for Online-Only Students 2003–2005*

AP exam	AP exams lifetime					Online AP courses 2003–2005				
	Mean	Std. dev.	Median	Mode	Max.	Mean	Std. dev.	Median	Mode	Max.
U.S. History	2.1	1.5	2	1	14	1.2	0.5	1	1	8
Microeconomics	4.6	3.2	4	1	17	1.6	0.7	1	1	6
Macroeconomics	5.3	3.1	5	3	17	1.6	0.8	1	1	6
English Language & Composition	2.4	1.9	2	1	15	1.2	0.6	1	1	6
English Literature & Composition	2.8	2.3	2	1	15	1.3	0.7	1	1	6
U.S. Government & Politics	4.0	2.8	3	1	17	1.4	0.7	1	1	6
Calculus AB	3.1	2.0	3	1	12	1.2	0.6	1	1	5
Psychology	2.8	2.2	2	1	17	1.2	0.6	1	1	6
Statistics	4.2	2.8	3	2	17	1.2	0.6	1	1	8



**Table B3***Sex and Ethnicity by AP Exam for Online-Only Students 2003–2005*

AP exam	% Female	% Am. Indian or Alaskan Native	% Black or African American	% Mexican- American	% Asian American or Pac. Islander	% Puerto Rican	% Other Hispanic	% White	% Other
U.S. History	52	1	5	5	9	1	3	73	3
Microeconomics	44	1	4	3	16	0	4	69	4
Macroeconomics	43	0	3	2	17	1	6	67	4
English Language & Composition	66	2	8	5	8	1	2	71	4
English Literature & Composition	63	1	11	3	5	1	2	73	3
U.S. Government & Politics	52	1	7	4	10	0	5	68	5
Calculus AB	41	2	6	3	15	0	4	67	3
Psychology	71	0	6	2	6	0	1	81	3
Statistics	40	1	3	3	18	0	1	70	4

**Table B4***Education Level by AP Exam for Online-Only Students 2003–2005*

AP exam	Freshmen		Sophomores		Juniors		Seniors	
	<i>N</i>	Row %	<i>N</i>	Row %	<i>N</i>	Row %	<i>N</i>	Row %
U.S. History	9	1	125	11	729	67	226	21
Microeconomics	14	2	66	8	235	27	557	64
Macroeconomics	6	1	54	6	209	22	689	72
English Language & Composition	7	1	44	5	528	62	271	32
English Literature & Composition	3	0	7	1	147	16	778	83
U.S. Government & Politics	17	1	59	5	286	22	923	72
Calculus AB	2	0	12	2	91	16	450	81
Psychology	6	0	73	6	395	31	795	63
Statistics	7	1	30	6	99	19	388	74

**Table B5***Comparison of 2005 No-Online and Online-Only Students by Sex by AP Exam*

AP exam	No online		Online only	
	Female	Male	Female	Male
	Row %	Row %	Row %	Row %
U.S. History*	56	44	49	51
Microeconomics	43	57	43	57
Macroeconomics	46	54	41	59
Eng. Lang & Composition	64	36	68	32
Eng. Lit. & Composition	64	36	63	37
U.S. Government & Politics	53	47	51	49
Calculus AB*	48	52	39	61
Psychology	66	34	69	31
Statistics*	51	49	38	62

\* Chi-sq statistic  $p$ -value < .05.

**Table B6**

*Comparison of 2005 No-Online and Online-Only Students by Ethnicity by AP Exam*

Student	AP exam	Am. Indian/ Alaskan Native	Black or African-American	Mexican- American	Asian American or Pac. Islander	Puerto Rican	Other Hispanic	White	Other
		Row %	Row %	Row %	Row %	Row %	Row %	Row %	Row %
No online	U.S. History*	0	6	5	12	1	4	67	4
Online only	U.S. History*	0	5	2	9	0	2	79	3
No online	Microeconomics	0	4	4	19	1	4	64	3
Online only	Microeconomics	0	3	1	20	1	3	70	3
No online	Macroeconomics*	0	4	6	19	1	5	62	3
Online only	Macroeconomics*	0	5	1	18	1	4	68	5
No online	Eng. Lang & Composition*	1	7	7	12	1	5	65	4
Online only	Eng. Lang & Composition*	3	12	3	8	0	3	69	3
No online	Eng. Lit. & Composition*	1	6	5	10	1	4	70	3
Online only	Eng. Lit. & Composition*	1	16	1	7	1	3	68	3
No online	U.S. Government & Politics*	0	5	6	13	1	5	66	4
Online only	U.S. Government & Politics*	0	7	3	11	0	6	69	3
No online	Calculus AB*	0	4	4	16	1	3	68	3
Online only	Calculus AB*	1	6	3	17	0	9	61	3
No online	Psychology*	0	6	3	13	1	4	69	4
Online only	Psychology*	0	6	1	7	0	1	81	3
No online	Statistics	0	4	3	17	1	3	68	3
Online only	Statistics	1	2	2	18	0	1	72	5

\* Chi-sq statistic  $p$ -value < .05.

**Table B7***Comparison of 2005 No-Online and Online-Only Students by Education Level by AP Exam*

AP Exam	No online				Online only			
	Freshmen	Sophomores	Juniors	Seniors	Freshmen	Sophomores	Juniors	Seniors
	Row %	Row %	Row %	Row %	Row %	Row %	Row %	Row %
U.S. History*	0	7	87	6	1	13	63	23
Microeconomics*	0	2	12	85	2	10	27	61
Macroeconomics*	0	1	9	90	1	5	21	73
Eng. Lang & Composition*	0	1	81	17	1	8	61	30
Eng. Lit. & Composition*	0	0	7	93	1	1	16	82
U.S. Government & Politics*	0	5	10	84	1	6	24	70
Calculus AB*	0	1	18	80	1	3	16	80
Psychology*	0	3	32	66	1	7	33	60
Statistics*	0	4	21	75	2	4	16	79

\* Chi-sq statistic  $p$ -value  $<.05$ .

**Appendix C**  
**PSAT/NMSQT Tables on Group Work**

**Table C1**

*Group-Work Regression for English Literature and Composition Free-Response Section  
Score With PSAT/NMSQT Flag Only*

Variable	Parameter estimate	Standard error	Type II S	F value	Pr > F	Standardized estimate
Intercept	32.46	1.38	80312.00	557.09	<.0001	0.00
Group demographics			17886.00	15.51	<.0001	
FEMALE	3.37	0.85	2282.89	15.84	<.0001	0.12
AF_AM	-12.56	1.36	12362.00	85.75	<.0001	-0.28
ASIAN	1.99	1.80	176.27	1.22	0.27	0.03
HISP	-7.10	1.68	2573.09	17.85	<.0001	-0.13
OTHR	-5.10	1.90	1036.37	7.19	0.01	-0.08
FR	6.39	6.98	120.68	0.84	0.36	0.03
SO	-2.03	4.61	27.91	0.19	0.66	-0.01
JR	1.27	1.13	183.10	1.27	0.26	0.03
Group academic history			8920.89	20.63	<.0001	
PSAT2	3.60	0.97	2003.50	13.90	0.00	0.12
EXAMS_LIFETIME	1.01	0.21	3284.06	22.78	<.0001	0.17
ONLINE_LIFETIME	0.80	0.62	235.48	1.63	0.20	0.04
Group hours online instruction			188.83	1.31	0.25	
HRSONLINE_GE1	1.52	1.33	188.83	1.31	0.25	0.03
Grouped group work			2985.10	5.18	0.00	
A few times a year	-1.96	1.04	511.54	3.55	0.06	-0.06
Once/twice a month	-3.35	1.35	892.63	6.19	0.01	-0.08
Once/twice a week	-4.64	1.28	1912.91	13.27	0.00	-0.11
About every day	-4.53	1.65	1092.86	7.58	0.01	-0.08

*Note:* Table reports regression statistics for hierarchically grouped and individual independent variables used in model. Overall  $R^2 = .22$  Grouped group-work variables partial  $R^2 = .02$ .

**Table C2*****Group-Work Regression for English Literature and Composition Multiple-Choice Section  
Score With PSAT/NMSQT Flag Only***

Variable	Parameter estimate	Standard error	Type II S	F value	Pr > F	Standardized estimate
Intercept	28.72	1.50	62878.00	364.71	<.0001	0.00
Group demographics			17664.00	12.81	<.0001	
FEMALE	0.39	0.93	30.67	0.18	0.67	0.01
AF_AM	-11.95	1.48	11202.00	64.97	<.0001	-0.24
ASIAN	-3.52	1.97	547.39	3.18	0.08	-0.05
HISP	-9.13	1.84	4253.32	24.67	<.0001	-0.15
OTHR	-8.07	2.08	2599.13	15.08	0.00	-0.11
FR	12.21	7.64	440.29	2.55	0.11	0.05
SO	6.91	5.04	324.06	1.88	0.17	0.04
JR	1.36	1.24	207.58	1.20	0.27	0.03
Group academic history			16669.00	32.23	<.0001	
PSAT2	4.78	1.06	3527.35	20.46	<.0001	0.14
EXAMS_LIFETIME	1.16	0.23	4277.54	24.81	<.0001	0.17
ONLINE_LIFETIME	2.17	0.68	1750.98	10.16	0.00	0.10
Group hrs online instruction			78.39	0.45	0.50	
HRSONLINE_GE1	0.98	1.45	78.39	0.45	0.50	0.02
Grouped group-work			8092.18	11.73	<.0001	
A few times a year	-3.95	1.14	2075.50	12.04	0.00	-0.11
Once/twice a month	-5.55	1.47	2446.56	14.19	0.00	-0.11
Once/twice a week	-6.99	1.39	4332.31	25.13	<.0001	-0.15
About every day	-8.14	1.80	3527.13	20.46	<.0001	-0.14

*Note:* Table reports regression statistics for hierarchically grouped and individual independent variables used in model. Overall  $R^2 = .28$ . Grouped group-work variables partial  $R^2 = .04$ .

**Table C3*****Group-Work Regression for English Literature and Composition Composite Score With PSAT/NMSQT Flag Only***

Variable	Parameter estimate	Standard error	Type II SS	F value	Pr > F	Standardized estimate
Intercept	61.18	2.53	285314.00	584.82	<.0001	0.00
Group demographics			67023.00	17.17	<.0001	
FEMALE	3.76	1.56	2842.81	5.83	0.02	0.07
AF_AM	-24.51	2.49	47100.00	96.54	<.0001	-0.29
ASIAN	-1.52	3.32	102.41	0.21	0.65	-0.01
HISP	-16.24	3.09	13443.00	27.55	<.0001	-0.15
OTHR	-13.17	3.50	6917.96	14.18	0.00	-0.11
FR	18.60	12.85	1021.99	2.09	0.15	0.04
SO	4.88	8.48	161.77	0.33	0.56	0.02
JR	2.63	2.08	780.60	1.60	0.21	0.04
Group academic history			49637.00	33.91	<.0001	
PSAT2	8.38	1.78	10848.00	22.23	<.0001	0.14
EXAMS_LIFETIME	2.17	0.39	15058.00	30.86	<.0001	0.19
ONLINE_LIFETIME	2.97	1.15	3270.69	6.70	0.01	0.08
Group hrs online instruction			510.55	1.05	0.31	
HRSONLINE_GE1	2.50	2.45	510.55	1.05	0.31	0.03
Grouped group work			20809.00	10.66	<.0001	
A few times a year	-5.92	1.92	4647.79	9.53	0.00	-0.09
Once/twice a month	-8.91	2.48	6294.81	12.90	0.00	-0.11
Once/twice a week	-11.64	2.35	12003.00	24.60	<.0001	-0.15
About every day	-12.68	3.03	8546.65	17.52	<.0001	-0.12

*Note:* Table reports regression statistics for hierarchically grouped and individual independent variables used in model. Overall  $R^2 = .28$ . Grouped group-work variables partial  $R^2 = .03$ .



**Table C4*****Group-Work Regression for Calculus AB Free-Response Section Score With PSAT/NMSQT Flag Only***

Variable	Parameter estimate	Standard error	Type II SS	F Value	Pr > F	Standardized estimate
Intercept	19.36	2.04	13952.00	90.47	<.0001	0.00
Group demographics			8297.30	6.73	<.0001	
FEMALE	-4.92	1.13	2936.66	19.04	<.0001	-0.18
AF_AM	-6.01	2.42	953.01	6.18	0.01	-0.10
ASIAN	2.05	1.64	241.29	1.56	0.21	0.05
HISP	-7.26	2.32	1511.99	9.80	0.00	-0.13
OTHR	-3.20	2.55	242.15	1.57	0.21	-0.05
FR	8.80	8.99	147.63	0.96	0.33	0.04
SO	9.54	3.88	933.75	6.05	0.01	0.10
JR	1.51	1.48	159.87	1.04	0.31	0.04
Group academic history			567.92	1.23	0.30	
PSAT2	-0.63	1.53	26.15	0.17	0.68	-0.02
EXAMS_LIFETIME	0.32	0.30	183.48	1.19	0.28	0.05
ONLINE_LIFETIME	1.12	0.91	230.64	1.50	0.22	0.05
Group hrs online instruction			777.26	5.04	0.03	
HRSONLINE_GE1	-4.08	1.82	777.26	5.04	0.03	-0.09
Group group-work			4708.44	7.63	<.0001	
A few times a year	-4.67	1.47	1552.19	10.06	0.00	-0.14
Once/twice a month	-2.92	2.15	284.29	1.84	0.18	-0.06
Once/twice a week	-0.24	2.04	2.08	0.01	0.91	-0.01
About every day	-8.00	1.58	3962.70	25.70	<.0001	-0.22

*Note:* Table reports regression statistics for hierarchically grouped and individual independent variables used in model. Overall  $R^2 = .17$ . Grouped group-work variables partial  $R^2 = .05$ .

**Table C5*****Group-Work Regression for Calculus AB Multiple-Choice Section Score With PSAT/NMSQT Flag Only***

Variable	Parameter estimate	Standard error	Type II SS	F value	Pr > F	Standardized estimate
Intercept	23.80	2.14	21081.00	123.71	<.0001	0.00
Group demographics			15747.00	11.55	<.0001	
FEMALE	-7.20	1.19	6292.03	36.92	<.0001	-0.24
AF_AM	-8.18	2.54	1765.09	10.36	0.00	-0.13
ASIAN	2.62	1.73	393.75	2.31	0.13	0.06
HISP	-8.17	2.44	1912.02	11.22	0.00	-0.13
OTHR	-0.79	2.68	14.80	0.09	0.77	-0.01
FR	3.91	9.45	29.23	0.17	0.68	0.02
SO	15.61	4.07	2502.28	14.68	0.00	0.15
JR	3.47	1.56	844.72	4.96	0.03	0.09
Group academic history			418.18	0.82	0.48	
PSAT2	0.71	1.61	33.00	0.19	0.66	0.02
EXAMS_LIFETIME	0.36	0.31	222.43	1.31	0.25	0.05
ONLINE_LIFETIME	0.42	0.96	32.92	0.19	0.66	0.02
Group hrs online instruction			1752.62	10.28	0.00	
HRSONLINE_GE1	-6.13	1.91	1752.62	10.28	0.00	-0.13
Grouped group-work			6999.98	10.27	<.0001	
A few times a year	-5.44	1.55	2101.95	12.33	0.00	-0.15
Once/twice a month	-4.23	2.26	596.32	3.50	0.06	-0.08
Once/twice a week	-2.74	2.14	279.54	1.64	0.20	-0.05
About every day	-10.20	1.66	6435.08	37.76	<.0001	-0.26

*Note:* Table reports regression statistics for hierarchically grouped and individual independent variables used in model. Overall  $R^2 = .24$ . Grouped group-work variables partial  $R^2 = .06$ .

**Table C6*****Group-Work Regression for Calculus AB Composite Score With PSAT/NMSQT Flag Only***

Variable	Parameter estimate	Standard error	Type II SS	F value	Pr > F	Standardized estimate
Intercept	43.16	4.02	69333.00	115.06	<.0001	0.00
Group demographics			46364.00	9.62	<.0001	
FEMALE	-12.12	2.23	17826.00	29.58	<.0001	-0.22
AF_AM	-14.19	4.78	5312.05	8.82	0.00	-0.12
ASIAN	4.68	3.25	1251.50	2.08	0.15	0.06
HISP	-15.43	4.59	6824.57	11.33	0.00	-0.14
OTHR	-3.99	5.05	376.66	0.63	0.43	-0.03
FR	12.71	17.77	308.24	0.51	0.47	0.03
SO	25.15	7.66	6493.15	10.78	0.00	0.13
JR	4.97	2.93	1739.54	2.89	0.09	0.07
Group academic history			1756.18	0.97	0.41	
PSAT2	0.08	3.03	0.40	0.00	0.98	0.00
EXAMS_LIFETIME	0.68	0.59	809.95	1.34	0.25	0.05
ONLINE_LIFETIME	1.54	1.81	437.83	0.73	0.39	0.04
Group hrs online instruction			4864.18	8.07	0.00	
HRSONLINE_GE1	-10.21	3.59	4864.18	8.07	0.00	-0.11
Grouped group-work			22971.00	9.53	<.0001	
A few times a year	-10.11	2.91	7266.68	12.06	0.00	-0.15
Once/twice a month	-7.15	4.25	1704.09	2.83	0.09	-0.07
Once/twice a week	-2.98	4.03	329.85	0.55	0.46	-0.03
About every day	-18.20	3.12	20497.00	34.02	<.0001	-0.25

*Note:* Table reports regression statistics for hierarchically grouped and individual independent variables used in model. Overall  $R^2 = .21$ . Grouped group-work variables partial  $R^2 = .06$ .

**Table C7*****Group-Work Regression for Statistics Free-Response Section Score With PSAT/NMSQT Flag***

Variable	Parameter estimate	Standard error	Type II SS	F value	Pr > F	Standardized estimate
Intercept	12.83	1.74	5063.64	54.18	<.0001	0.00
Group demographics			7377.04	9.87	<.0001	
FEMALE	-2.55	0.88	788.34	8.44	0.00	-0.11
AF_AM	-14.12	2.67	2617.32	28.01	<.0001	-0.21
ASIAN	-1.96	1.16	268.31	2.87	0.09	-0.07
HISP	-7.36	2.19	1054.45	11.28	0.00	-0.13
OTHR	-6.04	2.07	794.53	8.50	0.00	-0.11
FR	5.55	3.82	196.82	2.11	0.15	0.06
SO	6.70	1.92	1137.71	12.17	0.00	0.14
JR	3.14	1.13	720.14	7.71	0.01	0.11
Group academic history			7145.06	25.49	<.0001	
PSAT2	3.39	1.43	524.54	5.61	0.02	0.09
EXAMS_LIFETIME	1.11	0.17	4071.83	43.57	<.0001	0.28
ONLINE_LIFETIME	1.89	0.69	697.07	7.46	0.01	0.11
Group hrs online instruction			556.61	5.96	0.02	
HRSONLINE_GE1	-4.18	1.71	556.61	5.96	0.02	-0.09
Grouped group work			2112.50	5.65	0.00	
A few times a year	-3.42	1.18	783.58	8.38	0.00	-0.12
Once/twice a month	-4.58	1.59	781.13	8.36	0.00	-0.11
Once/twice a week	-5.70	1.55	1263.58	13.52	0.00	-0.15
About every day	-1.66	1.39	132.39	1.42	0.23	-0.05

*Note:* Table reports regression statistics for hierarchically grouped and individual independent variables used in model. Overall  $R^2 = .26$ . Grouped group-work variables partial  $R^2 = .03$ .

**Table C8*****Group-Work Regression for Statistics Multiple-Choice Section Score With PSAT/NMSQT Flag***

Variable	Parameter estimate	Standard error	Type II SS	F value	Pr > F	Standardized estimate
Intercept	19.71	1.85	11956.00	113.85	<.0001	0.00
Group demographics			9371.46	11.16	<.0001	
FEMALE	-3.97	0.93	1911.45	18.20	<.0001	-0.17
AF_AM	-12.27	2.83	1976.95	18.83	<.0001	-0.17
ASIAN	-2.82	1.23	554.40	5.28	0.02	-0.09
HISP	-7.17	2.32	1000.37	9.53	0.00	-0.12
OTHR	-4.14	2.19	373.65	3.56	0.06	-0.07
FR	7.26	4.05	336.92	3.21	0.07	0.07
SO	8.37	2.04	1775.23	16.91	<.0001	0.17
JR	4.76	1.20	1651.50	15.73	<.0001	0.16
Group academic history			6817.79	21.64	<.0001	
PSAT2	4.04	1.52	746.11	7.11	0.01	0.11
EXAMS_LIFETIME	1.17	0.18	4511.23	42.96	<.0001	0.28
ONLINE_LIFETIME	0.81	0.73	129.74	1.24	0.27	0.05
Group hrs online instruction			386.48	3.68	0.06	
HRSONLINE_GE1	-3.49	1.82	386.48	3.68	0.06	-0.07
Grouped group work			2105.34	5.01	0.00	
A few times a year	-3.75	1.25	945.24	9.00	0.00	-0.12
Once/twice a month	-4.19	1.68	651.94	6.21	0.01	-0.10
Once/twice a week	-5.19	1.64	1046.79	9.97	0.00	-0.13
About every day	-3.40	1.47	557.11	5.31	0.02	-0.09

*Note:* Table reports regression statistics for hierarchically grouped and individual independent variables used in model. Overall  $R^2 = .25$ . Grouped group-work variables partial  $R^2 = .03$ .

**Table C9*****Group-Work Regression for Statistics Composite Score With PSAT/NMSQT Flag***

Variable	Parameter estimate	Standard error	Type II SS	F value	Pr > F	Standardized estimate
Intercept	32.54	3.36	32581.00	93.75	<.0001	0.00
Group demographics			32820.00	11.80	<.0001	
FEMALE	-6.53	1.69	5154.91	14.83	0.00	-0.15
AF_AM	-26.39	5.15	9143.69	26.31	<.0001	-0.20
ASIAN	-4.78	2.23	1594.11	4.59	0.03	-0.08
HISP	-14.54	4.23	4108.93	11.82	0.00	-0.13
OTHR	-10.17	3.99	2257.89	6.50	0.01	-0.10
FR	12.80	7.37	1048.79	3.02	0.08	0.07
SO	15.07	3.70	5755.31	16.56	<.0001	0.16
JR	7.91	2.18	4552.77	13.10	0.00	0.14
Group academic history			27674.00	26.54	<.0001	
PSAT2	7.43	2.76	2521.81	7.26	0.01	0.11
EXAMS_LIFETIME	2.27	0.32	17155.00	49.36	<.0001	0.30
ONLINE_LIFETIME	2.70	1.33	1428.27	4.11	0.04	0.08
Group hrs online instruction			1870.73	5.38	0.02	
HRSONLINE_GE1	-7.67	3.31	1870.73	5.38	0.02	-0.09
Grouped group work			8243.74	5.93	0.00	
A few times a year	-7.17	2.28	3450.11	9.93	0.00	-0.13
Once/twice a month	-8.77	3.06	2860.31	8.23	0.00	-0.11
Once/twice a week	-10.90	2.99	4610.55	13.27	0.00	-0.15
About every day	-5.05	2.68	1232.65	3.55	0.06	-0.07

*Note:* Table reports regression statistics for hierarchically grouped and individual independent variables used in model. Overall  $R^2 = .28$ . Grouped group-work variables partial  $R^2 = .03$ .