Research Report No. 2002-9

An Investigation of the Validity of AP® Grades of 3 and a Comparison of AP and Non-AP Student Groups

Barbara G. Dodd, Steven J. Fitzpatrick, R. J. De Ayala, and Judith A. Jennings

College Board Research Report No. 2002-9

An Investigation of the Validity of AP® Grades of 3 and a Comparison of AP and Non-AP Student Groups

Barbara G. Dodd, Steven J. Fitzpatrick, R. J. De Ayala, and Judith A. Jennings

College Entrance Examination Board, New York, 2002

Barbara G. Dodd is a Professor of Educational Psychology at the University of Texas at Austin.

Steven J. Fitzpatrick is an Associate Director of the Measurement and Evaluation Center at the University of Texas at Austin.

R. J. De Ayala is a Professor of Educational Psychology at the University of Nebraska.

Judith A. Jennings is a Graduate Research Assistant at the University of Texas at Austin.

The College Board: Expanding College Opportunity

The College Board is a national nonprofit membership association whose mission is to prepare, inspire, and connect students to college and opportunity. Founded in 1900, the association is composed of more than 4,200 schools, colleges, universities, and other educational organizations. Each year, the College Board serves over three million students and their parents, 22,000 high schools, and 3,500 colleges through major programs and services in college admissions, guidance, assessment, financial aid, enrollment, and teaching and learning. Among its best-known programs are the SAT®, the PSAT/NMSQT®, and the Advanced Placement Program® (AP®). The College Board is committed to the principles of excellence and equity, and that commitment is embodied in all of its programs, services, activities, and concerns.

For further information, visit www.collegeboard.com.

Additional copies of this report (item #995384) may be obtained from College Board Publications, Box 886, New York, NY 10101-0886, 800 323-7155. The price is $15. Please include $4 for postage and handling.

Printed in the United States of America.

Contents

Abstract

The purpose of this study was to address the validity of grades of 3 on AP® Examinations and to compare AP students to other relevant student groups. While research has shown that students who earn grades of 3 or higher and place out of introductory courses do well in the subsequent courses, there are some college faculty members who think this is not always the case. In order to address this issue a number of different statistical techniques were employed to determine if finer gradations of the grade group of 3s might prove useful for course placement in college. More specifically, three different statistical techniques were used to identify "low" and "high" 3 score groups. These procedures provided insight to whether the "high" and "low" 3 groups are more similar to other 3-grade groups or the adjacent score groups (2s and 4s). The grade groups derived from the different statistical techniques were compared in terms of a number of college achievement variables such as grades in the subsequent course, number of college hours taken in the content area, and the grades earned in those courses. In addition, the performance of AP students who placed out of the first course in the sequence was compared to the performance of a matched non-AP group of students, AP students who took the introductory course prior to the subsequent course, and students who were concurrently enrolled in an equivalent college course while still in high school. The study was replicated for four different entering classes at a large university. Collectively, the findings of this study did not support finer gradations of the AP score category of 3. It was also found that AP students who earn credit by examination tend to make the same or higher grades in subsequent courses than do the other comparison groups.

I. Introduction

The College Board Advanced Placement Program® (AP) offers high school students the opportunity to take college-level classes while they are still attending high school. These courses are based on standard college curricula in 33 subject areas. During the month of May, AP students can take AP Examinations in an attempt to earn college course credit or advanced placement, or both, in a college course sequence.

Each AP Examination is composed of a multiple-choice section and a free-response section. The scores from the two sections are combined to form a composite score that is transformed onto a 1 to 5 scale. The College Board provides the following interpretation of the 1 to 5 scale:

5 – extremely well qualified
4 – well qualified
3 – qualified
2 – possibly qualified
1 – no recommendation

Research studies by Casserly (1986), Koch, Fitzpatrick, Triscari, Mahoney, and Cope (1988), Morgan and Crone (1993), Morgan and Ramist (1998), and Morgan and Maneckshana (2000) have found that AP students who place out of introductory college courses and into the second college course in a sequence of courses perform as well as if not better than students who took the prerequisite college course. Several of these studies (Koch et al., 1988; Morgan and Maneckshana, 2000) also found that students who took AP courses and examinations in high school tended to take more courses in the AP subject area than students who did not take AP courses.

Despite the fact that these studies have shown AP students perform as well if not better than other groups of college students, some college faculty members report that these overall findings are not always the case. Some faculty members think that there are students who receive advanced standing who are placed too high. It is possible that the 3-grade group includes a wide range of ability levels that leads to these apparently disparate points of view.

Thus, the purpose of the present research was to further examine the performance of AP students with grades of 3 who are placed into sequent courses and determine if finer gradations of the grade group of 3s is warranted. In addition, the performance of AP students who were placed into the sequent course was compared to the performance of several other groups of students.

II. Research Questions

Questions 1–3 concern finer gradations of the 3-grade group. Questions 4–6 address the performance of AP students relative to other student groups.

1. On what basis would AP distinguish between low and high 3s?
2. How would AP make the distinction between a low 3 and a high 3?
3. Given the performance in college of those with low and high AP grades of 3, should AP distinguish students receiving low and high grades of 3?
4. How do AP students who took the introductory college course before the sequent course compare to AP students who placed out of the introductory course?
5. How do non-AP students who took the introductory college course before the sequent course compare to AP students who place out of the introductory course?

6. How do the AP students who placed out the introductory course compare to students who were concurrently enrolled in the introductory college level course while still in high school?

III. Design and Methodology

Data Source

AP Examinations in Calculus AB, English Language and Composition, and Biology have historically had some of the largest numbers of examinees. For each of the entering first year classes at the University of Texas at Austin from 1996–1999, 956 to 1,181 students have Calculus AB scores, 1,235 to 1,838 students have English Language and Composition scores, and 430 to 529 students have Biology scores. Each of these AP Examinations allows students to place out of part of a two- or three-course sequence. A score of 3 or higher on the AP Calculus AB Exam yields credit by examination for Mathematics 408C, the first course in a two-course sequence in college calculus. Similarly, a score of 3 or higher on the AP English Language and Composition Examination yields credit for the Rhetoric and Composition course—E 306. Students typically take an English literature course (E 316K) as the next course at the University of Texas at Austin. The AP Examination in Biology is a bit different from the previously mentioned examinations in that a score of 5 yields credit for the entire sequence of introductory courses (BIO 302, 303, and 304). A score of 3 or 4 yields credit for two of the three-course sequence (BIO 302, 304).

In order to generalize the study's findings, four different entering classes (1996–1999) were used in the present research. A preliminary analysis of these four entering classes revealed that these students took the majority of their AP Examinations in the selected subject areas in the 1995–1999 test administrations. Thus the data for this research included five different test administration dates for the four entering classes.

Educational Testing Service (ETS) provided item response vectors, section scores (multiple-choice and constructed response), and composite scores on the AP Examinations for five test dates for both national samples and the four University of Texas at Austin samples. For the University of Texas at Austin samples, ETS was provided with a file containing the necessary identifying information for the freshmen who sent AP scores. Once the University of Texas at Austin data sets had been received from ETS, they were merged with additional data obtained from the University of Texas at Austin student record database. Variables obtained from the University of Texas at Austin database included grades in college courses taken in the subjects of the AP Examinations, number of hours of college courses taken in the subjects of the AP Examinations, high school rank, and admissions test scores (SAT® and ACT). In addition, ETS provided the composite score cut points for each of the test dates for each examination so that the possible range of composite scores for each AP grade group could be used in the analyses.

High school grades in AP courses were not used in the analyses because some courses listed as AP classes on high school transcripts received by the University of Texas at Austin are not College Board AP courses. Also, some high schools award additional grade points to AP classes, while others do not. Thus, the decision was made to not include high school grades, but rather focus the research on college course work where there is more control over the quality of the data.

The University of Texas at Austin AP samples were subdivided into three AP groups:

1. AP students who received credit by examination in the entire course sequence.
2. AP students who received credit by examination for at least one course but not all courses in the course sequence.
3. AP students who received no credit by examination for any of the courses in the course sequence.

Two additional comparison groups were also identified for the University of Texas at Austin samples:

4. A non-AP group that was matched to the second subgroup of the AP students listed above using high school rank and SAT Total scores. For students who had only ACT scores, the ACT Composite scores were converted to SAT Total scores using the College Board conversion tables (Dorans, 1999).
5. Students who were concurrently enrolled in a college course that is considered to be equivalent to the first course in the sequence while they were still in high school.

Data Analyses

The data analyses can be broken into two phases. The first phase addresses research questions 1–3, while the second phase concerns research questions 4–6.

Phase I. Research questions 1–3 were addressed using a two-stage approach. The first stage examines

different approaches for defining "high" and "low" grades of 3. Because the section scores and item response vectors are not equated across test years, the analyses were performed separately for each test administration year. Three different techniques were investigated.

1. "High" and "low" grades of 3 were created using a series of fixed percentages of the examinees from the top and bottom of the composite score distribution using the national sample for each testing year. Two different sets of percentages were used to divide the grade group of 3s. For one method, the 3-grade group was divided into thirds to form the high, middle, and low 3 groups; the same procedure was followed to divide the 2- and 4-grade groups into high, middle, and low groups so that they could be compared to the newly formed groups of 3 scores. For the other fixed percentage method, the 3-grade group was divided into top 25 percent, middle 50 percent, and bottom 25 percent; the 2- and 4-grade groups were also subdivided in a similar fashion in order to form low, middle, and high subgroups for comparison purposes. The obtained cut points from the national samples were then applied to the University of Texas at Austin samples for use in the second stage of Phase I.

2. Using the national samples, k-means cluster analyses of section scores and response vectors were used to identify clusters of examinees. The scores on the multiple-choice and constructed response sections were standardized within a given test administration year so that the differences in the scales for the two section scores would not give greater weight to the section score with more score points. The item level data were also standardized for the same reason. For the section score and item level analyses, two different clusters analyses that varied in terms of the number of clusters specified were conducted using each of the national data sets.

   The numbers of clusters investigated were seven and eleven. It was thought that seven clusters might represent a subdividing of the 3-grade group into three groups while leaving the other four grade groups (1, 2, 4, and 5) intact. The eleven clusters on the other hand might represent dividing the 2s, 3s, and 4s into three subgroups each, while leaving the 1s and 5s intact as a single group, respectively. Unlike the fixed percentage methods, all AP scores were used because it was not clear how the cluster analyses would treat the scores of 1 and 5.

   The clusters of examinees obtained from each cluster analysis were then related to the five AP score categories of performance to determine the nature of the clusters. Finally, using the cluster centers obtained from each of the cluster analyses of the national samples, individuals in the University of Texas at Austin samples were assigned to clusters. The cluster membership was then used in the second stage of Phase I.

3 A latent class analysis (LCA) was performed to examine whether examinees with a grade of 3 should or should not be considered a homogeneous group. For this analysis it was hypothesized that the examinees with AP grades of 3 consist of a mixture of latent populations or classes. LCA was performed on each of the five test administrations (1995–1999) for each of the three examinations. Using the examinees' response vectors, the best fitting and interpretable latent class model (LCM) was determined. It was believed that the best fitting LCM would consist of two or at most three classes. For a three-class situation, it was expected that the LCM's interpretation would identify one class as a high 3 group, a second class as a non-high and non-low 3 group, and the third class the low 3 group. Follow-up analyses on this latter class could identify the individuals who place out of the first semester course in the sequence but who do not perform well in the sequent course. A two-class solution might consist of high 3s in one class and low 3s in the other class. Upon determining an acceptable LCM each examinee's latent class membership was determined for use in the second stage of Phase I.

The second stage of Phase I examined the usefulness of the three strategies outlined in stage 1 by determining if the newly formed groups differ significantly on three academic outcome measures. These measures were a) grade in the sequent course, b) GPA for the courses taken in the AP subject area, and c) number of hours taken in the AP subject area. For each test administration year, an analysis of variance (ANOVA) was conducted to compare the groups on each of these outcome measures. Tukey's post hoc comparison test was used to determine which group means differed significantly from one another.

Phase II. Research questions 4, 5, and 6 involved the identification of the following four groups of students for each entering class:

1. AP students who received credit by examination for at least one course but not all courses in the course sequence.

2. AP students who received no credit by examination for any of the courses in the course sequence.

3. A non-AP group that was matched to the first subgroup of the AP students listed above using high school rank and SAT Total scores. The matching was accomplished by dividing high school rank into five categories and SAT Total scores into 100-point score intervals and then assigning the AP students to the appropriate subgroup based on their high school rank and SAT Total score. The University of Texas at Austin student records were then searched to identify enough non-AP students to equal the number of AP students in each of the subgroups. It should be noted that if a student had an ACT score rather than a SAT score, the ACT Composite score was converted to a SAT Total score using the tables of concordance provided by the College Board Report No. 99-1 (Dorans, 1999).

4. Students who were concurrently enrolled in a college course that is considered to be equivalent to the first course in the sequence while they were still in high school. These students earned college credit for the course while still in high school. There was no way to determine if they took the course on the college campus or the high school campus.

It should be noted that students who earned credit by examination for the entire course sequence via any number of different tests were not included in the Phase II analyses because they did not take the subsequent course. Thus, only students who actually took the subsequent course were included in these analyses.

The analyses consisted of descriptive statistics and ANOVAs. Comparison of the two AP groups answers research question 4 and comparison of the AP group 2 and the non-AP group addresses research question 5. Research question 6 is assessed by the comparison of the concurrently enrolled students with the AP students.

IV. Results

Description of Samples

The AP score distributions for each of the AP Examinations in English Language and Composition, Calculus AB, and Biology for the national samples by test administration year are presented in Table 1. These data were used to establish finer gradations of the AP score scale in the first phase of the study. More specifically they were used in the fixed percentage conditions and the cluster analyses conditions.

Table 2 presents the AP score frequency distributions for the three AP Examinations for the 1995–1999 test administrations for the University of Texas at Austin samples. These data were used to assess the adequacy of the finer gradations of the AP score scale that were identified in stage 1 of Phase I.

Table 3 shows the number and percentage of students in each of the University of Texas at Austin entering classes who either a) took the class that the AP Examination grants credit for at the University of Texas at Austin, b) were concurrently enrolled in an equivalent college course while in high school, c) took the course through correspondence, d) transferred course credit from another college, or e) earned credit by

TABLE 1

Frequency of AP® Scores by Test Administration Date for the National Data Set

		AP Examination					
Test Date	*AP Score*	*English Language and Composition*		*Calculus AB*		*Biology*	
		Freq.	*%*	*Freq.*	*%*	*Freq.*	*%*
1995	1	5,313	10.4	23,276	22.6	8,222	13.5
	2	18,926	36.9	18,950	18.4	13,933	22.8
	3	15,061	29.4	28,876	28.1	15,138	24.8
	4	8,471	16.5	19,590	19.1	13,269	21.7
	5	3,532	6.9	12,128	11.8	10,550	17.3
	Total	51,303		102,820		61,112	
1996	1	4,167	7.1	20,294	19.3	9,339	14.2
	2	18,308	31.0	21,082	20.1	14,340	21.9
	3	20,079	34.0	28,363	27.0	15,348	23.4
	4	12,244	20.8	21,871	20.8	14,301	21.8
	5	4,183	7.1	13,430	12.8	12,298	18.7
	Total	58,981		105,040		65,626	
1997	1	3,971	5.9	22,795	20.5	8,076	11.4
	2	19,372	28.9	22,122	19.9	14,830	21.0
	3	24,058	35.9	30,348	27.3	17,827	25.3
	4	12,536	18.7	22,686	20.4	15,752	22.3
	5	7,024	10.5	13,251	11.9	14,086	20.0
	Total	66,961		111,202		70,571	
1998	1	4,962	6.2	18,876	16.2	12,190	16.2
	2	22,889	28.8	20,732	17.8	17,271	22.9
	3	26,935	33.9	31,287	26.9	17,914	23.7
	4	17,152	21.6	27,103	23.3	13,893	18.4
	5	7,508	9.5	18,522	15.9	14,197	18.8
	Total	79,446		116,520		75,465	
1999	1	4,521	4.7	21,634	17.2	10,638	13.1
	2	31,971	33.1	24,000	19.1	17,766	21.8
	3	33,861	35.0	31,339	24.9	19,147	23.5
	4	17,654	18.3	28,535	22.7	17,978	22.1
	5	8,696	9.0	20,271	16.1	15,963	19.6
	Total	96,703		125,779		81,492	

TABLE 2

Frequency of AP Scores by Test Administration Date for the University of Texas at Austin Data Set

Test Date	AP Score	AP Examination: English Language and Composition Freq.	%	Calculus AB Freq.	%	Biology Freq.	%
1995	1	21	2.2	4	5.2	9	5.3
	2	241	25.7	3	3.9	26	15.2
	3	362	38.6	11	14.3	40	23.4
	4	228	24.3	26	33.8	58	33.9
	5	87	9.3	33	42.9	38	22.2
	Total	939		77		171	
1996	1	25	1.7	93	9.9	28	6.4
	2	300	21.0	150	15.9	57	13.0
	3	573	40.1	266	28.3	114	26.0
	4	399	27.9	272	28.9	138	31.5
	5	133	9.3	160	17.0	101	23.1
	Total	1,430		941		438	
1997	1	21	1.4	126	11.8	32	6.4
	2	264	17.2	178	16.7	88	17.6
	3	683	44.6	329	30.9	130	26.1
	4	395	25.8	280	26.3	134	26.9
	5	169	11.0	152	14.3	115	23.0
	Total	1,532		1,065		499	
1998	1	25	1.3	120	11.8	46	9.4
	2	401	21.7	158	15.5	87	17.8
	3	748	40.4	294	28.8	135	27.7
	4	479	25.9	263	25.8	130	26.6
	5	199	10.7	186	18.2	90	18.4
	Total	1,852		1,021		488	
1999	1	9	2.7	139	12.9	28	8.2
	2	66	19.8	185	17.2	72	21.2
	3	145	43.4	280	26.0	82	24.1
	4	92	27.5	283	26.3	89	26.2
	5	22	6.6	189	17.6	69	20.3
	Total	334		1,076		340	

examination for the course. It should be noted that the credit by examination group includes not only AP students but also students who earned credit by examination via other examinations. These breakdowns are important because they are used to form the comparison groups in the second phase of the study.

Phase I: Fixed Percentages

The results for the fixed percentage procedures used to form the new AP score groups are presented separately for the two methods used. The results based on the middle-third method are presented first, followed by the results for the middle 50 percent method.

Middle-third method. Table 4 presents descriptive statistics of the grades in the subsequent English course, other English semester hours taken, and the grade point averages earned in the other English courses for each of the nine AP score groups in the University of Texas at Austin sample for each test date. Similar descriptive information is provided for AP Calculus score groups and the AP Biology score groups in Tables 5 and 6, respectively. While ANOVAs were run to determine statistically significant mean differences between the AP score groups on the various outcome measures, it was decided that the findings were not particularly meaningful given the small sample sizes. As a consequence, the data from the various test administrations were combined. This was possible for the current analyses because the AP score groups were determined from percentile ranks using the national samples. This is similar to equating the finer gradations of the AP 1–5 scale with equipercentile equating.

Table 7 shows the means, standard deviations, and frequencies for each of the AP score groups on each

TABLE 3

Description of the University of Texas at Austin Freshman Classes for 1996–1999

Course	1996 Freq.	%	1997 Freq.	%	1998 Freq.	%	1999 Freq.	%	Total Freq.	%
E 306										
Class	1,732	33.2	2,463	38.0	1,945	32.9	1,549	24.4	7,689	32.1
Concurrent	420	8.0	608	9.4	733	12.4	783	12.4	2,544	10.7
Correspondence	14	0.3	13	0.2	12	0.2	3	0.0	42	0.2
Transfer	700	13.4	745	11.5	577	9.8	393	6.2	2,415	10.1
CBE	1,951	37.4	2,148	33.2	2,086	35.3	2,532	40.0	8,717	36.4
None	402	7.7	502	7.7	556	9.4	1,075	17.0	2,535	10.6
M 408C										
Class	1,751	33.6	2,065	31.9	2,051	34.7	2,101	33.2	7,968	33.3
Concurrent	2	0.0	7	0.1	7	0.1	3	0.0	19	0.1
Correspondence	3	0.1	5	0.1	1	0.0	1	0.0	10	0.0
Transfer	0	0.0	0	0.0	1	0.0	0	0.0	1	0.0
CBE	843	16.1	870	13.4	925	15.7	1,000	15.8	3,638	15.2
None	2,620	50.2	3,532	54.5	2,924	49.5	3,230	51.0	12,306	51.4

Continued on page 6

TABLE 3 *Continued from page 5*

Description of the University of Texas at Austin Freshman Classes for 1996–1999

	1996		*1997*		*1998*		*1999*		*Total*	
Course	*Freq.*	*%*	*Freq.*	*%*	*Freq.*	*%*	*Freq.*	*%*	*Freq.*	*%*
BIO 302										
Class	979	18.8	1,151	17.8	927	15.7	588	9.3	3,645	15.2
Concurrent	58	1.1	63	1.0	57	1.0	69	1.1	247	1.0
Correspondence	1	0.0	0	0.0	0	0.0	0	0.0	1	0.0
Transfer	37	0.7	34	0.5	48	0.8	42	0.7	161	0.7
CBE	314	6.0	352	5.4	330	5.6	348	5.5	1,344	5.6
None	3,830	73.4	4,879	75.3	4,547	76.9	5,288	83.5	18,544	77.5
BIO 303										
Class	749	14.4	860	13.3	736	12.5	493	7.8	2,838	11.9
Concurrent	46	0.9	49	0.8	42	0.7	51	0.8	188	0.8
Correspondence	1	0.0	3	0.0	1	0.0	0	0.0	5	0.0
Transfer	20	0.4	27	0.4	32	0.5	19	0.3	98	0.4
CBE	127	2.4	141	2.2	106	1.8	135	2.1	509	2.1
None	4,276	81.9	5,399	83.3	4,992	84.5	5,637	89.0	20,304	84.8
BIO 304										
Class	599	11.5	734	11.3	525	8.9	493	7.8	2,351	9.8
Concurrent	2	0.0	2	0.0	1	0.0	0	0.0	5	0.0
Correspondence	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Transfer	21	0.4	8	0.1	5	0.1	2	0.0	36	0.2
CBE	315	6.0	359	5.5	345	5.8	367	5.8	1,386	5.8
None	4,282	82.1	5,376	83.0	5,033	85.2	5,473	86.4	20,164	84.2
Class Size	5,219		6,479		5,909		6,335		23,942	

of the three outcome measures for each of the AP Examinations. For the English AP Examination, there were statistically significant differences among the AP score groups in terms of the average grade earned in the sequent course ($F = 4.20$, $p < .0001$). Tukey post hoc comparison tests revealed the AP score groups of 4, 4-, 3+, and 3 earned significantly higher grades on average in the sequent course than did the 2- group. The AP score group 4 also earned a significantly higher average grade in the sequent course than did the 2+ AP score group. No other statistically significant differences were found. It should be noted that with the exception of the 2- AP group, all AP score groups earned average grades of B in the sequent course. In addition, none of the AP score groups differed significantly in the average numbers of other hours taken in English or the average GPAs in the other English courses.

The results for the Calculus AB Examination were similar to those found for the English Language and Composition Examination in that significant differences between the AP score groups were found only for the average grades earned in the sequent math course ($F = 3.23$, $p < .0001$). Tukey post hoc comparisons revealed that the 4+ and 4 AP groups earned significantly higher average grades than did the 2-, 2, and 3- AP score groups. The 4- AP group was significantly different from the 2- and 2 AP groups. No other differences were statistically significant.

The AP Examination in Biology also showed significant differences for the average grade in the next course in the sequence ($F = 3.23$, $p < .0014$). The AP 4+ group earned a significantly higher average grade in the sequent course than did the 3 or 3- AP score groups. In addition, there were significant differences in the average GPAs in other Biology courses ($F = 2.95$, $p < .0035$). Specifically, the 4+ AP score group had a significantly higher average GPA than did either the 3+ or 2 AP score groups. No other statistically significant differences were obtained for the analysis of the Biology outcome measures.

For all three of the AP Examinations, the 3- AP score group was not found to differ significantly from the 3 or 3+ AP score groups on any of the outcome measures. Nor did any of the AP groups of 3 differ significantly from the 2+ AP group or the 4- AP group.

Middle 50 percent method. Tables 8, 9, and 10 present the descriptive statistics of the outcome measures for each of the AP score groups by test administration year for the English Language and Composition Examination, Calculus AB Examination,

TABLE 4

Descriptive Statistics for the English Outcome Measures by Test Year and AP Score Groups Created by the Middle-Third Method

Test	*AP*	*E 316K Grades*			*Other English Hours Taken*			*GPAs in Other English Classes*		
Date	*Score*	*Mean*	*SD*	*Freq.*	*Mean*	*SD*	*Freq.*	*Mean*	*SD*	*Freq.*
1995	2-	3.07	0.90	28	7.33	7.52	9	3.10	0.79	9
	2	3.14	0.72	36	8.33	9.22	9	2.93	1.17	9
	2+	3.27	0.91	30	5.00	4.90	6	3.37	0.50	6
	3-	3.19	0.93	21	3.00	0.00	4	3.25	0.96	4
	3	3.24	0.72	25	3.00	0.00	5	3.20	0.45	5
	3+	3.39	0.85	18	7.80	9.15	5	3.58	0.53	5
	4-	3.40	0.89	5	3.00		1	3.00		1
	4	3.75	0.46	8	3.00	0.00	5	3.80	0.45	5
	4+	4.00	0.00	3	3.00		1	4.00		1
1996	2-	2.88	0.88	25	7.67	7.81	9	3.49	0.57	9
	2	3.17	0.81	36	3.86	1.46	7	3.36	0.85	7
	2+	3.12	0.73	49	6.92	6.64	13	3.25	1.12	13
	3-	3.08	0.84	61	5.75	7.17	12	3.28	0.44	12
	3	3.11	0.96	45	6.90	3.48	10	3.28	0.62	10
	3+	3.23	0.94	47	9.46	10.60	13	3.45	1.12	13
	4-	3.39	0.99	23	4.20	2.68	5	3.80	0.45	5
	4	3.58	0.76	26	6.33	5.50	9	3.29	1.30	9
	4+	3.33	0.65	12	3.00	0.00	2	4.00	0.00	2
1997	2-	2.68	1.04	22	6.60	2.51	5	3.50	0.37	5
	2	2.72	0.98	25	3.60	1.34	5	3.30	0.45	5
	2+	2.90	1.03	40	3.00	0.00	7	3.00	0.82	7
	3-	2.93	0.91	43	3.82	2.71	11	3.02	0.95	11
	3	3.26	0.71	53	5.25	6.90	12	3.47	0.50	12
	3+	3.19	1.00	37	7.07	7.31	14	3.69	0.52	14
	4-	3.33	1.02	21	6.40	5.42	8	3.40	0.89	8
	4	3.17	1.11	12	3.00		1	4.00		1
	4+	2.40	1.58	10	4.00	1.73	3	2.83	1.26	3
1998	2-	2.68	0.84	22	3.00	0.00	5	2.80	0.45	5
	2	3.19	0.56	27	3.00	0.00	4	3.50	0.58	4
	2+	2.75	1.13	16	3.00	0.00	3	3.00	1.00	3
	3-	3.33	0.73	27	3.75	1.50	4	3.50	1.00	4
	3	3.07	1.17	27	3.43	1.13	7	2.71	0.95	7
	3+	2.88	0.90	24	7.00	6.93	3	3.60	0.69	3
	4-	3.75	0.45	12	3.00		1	3.00		1
	4	3.38	0.74	8	3.00		1	4.00		1
	4+	3.17	0.75	6			0			0
1999	2-	2.43	1.27	7	3.00		1	3.00		1
	2	3.00	1.00	3			0			0
	2+			0			0			0
	3-	2.83	0.75	6	3.00	0.00	2	3.50	0.71	2
	3	3.00		1			0			0
	3+	3.29	0.76	7	3.00		1	4.00		1
	4-	4.00	0.00	2			0			0
	4	4.00		1			0			0
	4+	3.50	0.71	2	3.00		1	4.00		1

TABLE 5

Descriptive Statistics for the Mathematics Outcome Measures by Test Year and AP Score Groups Created by the Middle-Third Method

Test	*AP*	*M 408C Grades*			*Other Math Hours Taken*			*GPAs in Other Math Classes*		
Date	*Score*	*Mean*	*SD*	*Freq.*	*Mean*	*SD*	*Freq.*	*Mean*	*SD*	*Freq.*
1995	2-			0			0			0
	2	2.00		1	3.00	0.00	2	0.50	0.71	2
	2+			0			0			0
	3-	3.00		1	11.00		1	3.64		1
	3	1.67	1.53	3	4.00		1	0.00		1
	3+			0			0			0
	4-	2.00	1.41	2	7.00	0.00	2	2.29	1.62	2
	4	2.57	1.27	7	4.50	1.73	4	2.18	1.67	4
	4+	3.50	1.00	4	25.00	26.87	2	2.72	1.11	2
1996	2-	2.23	1.09	13	10.57	10.21	14	3.07	0.43	14
	2	2.09	1.38	11	6.71	2.21	7	2.61	0.96	7
	2+	2.75	0.91	20	6.40	3.94	15	2.65	0.75	15
	3-	2.57	1.28	37	7.52	5.57	23	2.88	0.68	23
	3	2.49	0.93	37	8.04	6.90	27	2.85	1.02	27
	3+	2.26	1.27	43	7.03	3.68	34	2.52	1.14	34
	4-	2.56	1.18	50	10.18	7.26	33	2.45	1.06	33
	4	2.85	1.23	54	9.49	7.01	39	2.90	0.98	39
	4+	2.78	1.29	60	8.02	6.28	43	2.94	1.10	43
1997	2-	2.11	1.41	19	4.67	1.87	9	3.00	0.71	9
	2	2.30	1.37	30	7.73	6.24	22	2.34	0.99	22
	2+	2.78	1.04	23	6.59	2.60	17	2.47	1.32	17
	3-	2.75	0.96	59	7.95	7.02	40	2.58	1.06	40
	3	2.63	1.13	57	8.26	6.22	42	2.88	0.90	42
	3+	2.79	1.01	67	8.68	4.92	47	2.64	1.05	47
	4-	3.06	0.95	51	6.19	3.29	36	2.90	0.93	36
	4	2.66	1.22	65	7.71	3.33	48	2.85	1.16	48
	4+	2.87	1.28	63	8.87	6.45	54	2.92	0.97	54
1998	2-	2.16	1.34	19	5.45	1.86	11	1.77	0.80	11
	2	2.24	1.36	25	5.86	1.79	14	2.74	0.99	14
	2+	2.15	1.46	20	6.82	3.52	11	2.06	1.03	11
	3-	2.30	1.17	43	6.42	4.05	33	2.55	0.95	33
	3	2.38	1.29	47	6.07	3.35	27	2.13	1.20	27
	3+	2.91	1.12	57	6.23	2.80	35	2.75	1.12	35
	4-	2.79	1.30	56	6.90	4.36	39	2.61	1.06	39
	4	3.12	0.96	50	7.44	3.60	34	2.92	0.85	34
	4+	2.90	1.13	68	6.78	3.41	50	2.68	1.10	50
1999	2-	2.47	1.26	19	4.07	1.33	14	2.95	0.82	14
	2	2.67	1.13	24	4.55	1.21	11	2.69	1.27	11
	2+	2.66	1.23	29	4.00	1.03	18	3.00	0.89	18
	3-	2.42	1.35	52	4.83	2.36	29	2.40	1.14	29
	3	2.94	1.11	49	5.06	2.26	33	2.67	1.07	33
	3+	2.73	1.34	40	5.79	3.14	19	2.67	1.10	19
	4-	3.07	1.16	43	5.58	2.09	31	2.66	1.06	31
	4	3.23	1.00	57	5.26	2.01	35	2.84	1.04	35
	4+	3.27	1.03	55	6.03	2.74	36	2.82	0.91	36

TABLE 6

Descriptive Statistics for the Biology Outcome Measures by Test Year and AP Score Groups Created by the Middle-Third Method

Test Date	*AP Score*	*BIO 303 Grades*			*Other Biology Hours Taken*			*GPAs in Other Biology Classes*		
		Mean	*SD*	*Freq.*	*Mean*	*SD*	*Freq.*	*Mean*	*SD*	*Freq.*
1995	2-	3.00		1	2.00		1	4.00		1
	2	3.50	0.71	2	2.00		1	4.00		1
	2+	2.33	1.53	3	3.50	2.12	2	3.00	1.41	2
	3-	2.33	1.15	3	4.00	2.83	2	3.08	0.12	2
	3	3.00	0.82	4	2.00	0.00	3	2.67	0.58	3
	3+	3.00	0.00	3	2.00	0.00	2	4.00	0.00	2
	4-	2.90	1.45	10	2.67	1.63	6	3.67	0.52	6
	4	2.40	0.84	10	3.00	1.73	3	3.33	0.58	3
	4+	3.86	0.38	7	2.33	0.58	3	3.67	0.58	3
1996	2-	1.67	0.58	3	2.00	0.00	2	2.50	0.71	2
	2	2.60	0.55	5	9.00	0.00	2	2.11	0.47	2
	2+	1.83	0.98	6	7.33	4.04	3	2.67	0.76	3
	3-	2.33	1.29	15	2.50	0.97	10	3.24	0.94	10
	3	2.50	1.26	16	2.70	2.21	10	2.70	1.06	10
	3+	3.14	0.94	22	3.87	2.47	15	2.78	1.08	15
	4-	3.13	0.96	16	3.33	2.00	9	3.00	0.73	9
	4	3.93	1.03	15	5.50	4.45	10	3.30	0.68	10
	4+	3.11	0.96	28	3.19	2.26	16	3.45	0.62	16
1997	2-	2.80	0.84	5	4.25	2.22	4	2.67	1.14	4
	2	2.40	1.17	10	5.29	2.75	7	2.11	1.26	7
	2+	3.50	1.00	4	5.00	4.08	4	2.83	0.97	4
	3-	2.54	1.13	13	6.25	5.15	8	2.84	0.59	8
	3	2.41	1.33	17	4.09	3.08	11	3.11	0.64	11
	3+	2.53	0.90	19	5.58	4.08	12	2.77	1.01	12
	4-	2.57	1.28	14	6.64	3.93	11	3.06	0.89	11
	4	2.79	1.27	19	7.09	4.16	11	2.95	1.17	11
	4+	3.33	1.11	15	5.00	3.46	7	3.23	0.69	7
1998	2-	2.80	1.64	5	5.20	3.56	5	2.82	0.63	5
	2	2.67	1.15	3	3.50	2.12	2	2.30	1.84	2
	2+	2.75	0.71	8	4.80	2.05	5	2.88	0.39	5
	3-	2.45	0.82	11	3.67	2.00	9	2.20	1.10	9
	3	2.87	0.83	15	4.30	1.77	10	2.83	1.12	10
	3+	2.50	1.45	14	4.17	3.56	12	2.62	1.37	12
	4-	3.12	0.93	17	4.13	2.72	15	2.98	0.74	15
	4	2.63	1.21	19	4.67	2.61	12	3.28	0.94	12
	4+	3.11	1.33	19	5.64	3.26	11	3.38	0.85	11
1999	2-	2.75	0.50	4	4.00	1.41	2	2.77	0.61	2
	2	2.33	0.58	3	4.00		1	2.00		1
	2+	2.33	2.08	3	3.00		1	2.33		1
	3-	2.67	0.58	3	2.00		1	2.00		1
	3	2.00	1.26	11	3.43	1.81	7	2.67	0.67	7
	3+	2.33	1.66	9	4.17	1.47	6	2.53	1.03	6
	4-	3.00	1.41	8	3.75	1.50	4	2.93	0.39	4
	4	3.11	0.78	9	4.00	1.93	8	2.52	1.23	8
	4+	3.30	0.67	10	3.25	0.71	8	3.32	1.38	8

TABLE 7

Descriptive Statistics for the Academic Outcome Measures Where the AP Score Groups Were Created by the Middle-Third Method

AP Exam	AP Score	Next Course Grades			Other Hours Taken in Subject Area			GPAs in Other Classes Taken in Subject Area		
		Mean	SD	Freq.	Mean	SD	Freq.	Mean	SD	Freq.
English Language and Composition	2-	2.82	0.94	104	6.41	6.14	29	3.23	0.63	29
	2	3.07	0.79	127	5.28	5.89	25	3.21	0.88	25
	2+	3.04	0.92	135	5.17	5.13	29	3.19	0.90	29
	3-	3.09	0.85	158	4.36	4.63	33	3.23	0.76	33
	3	3.18	0.88	151	5.03	4.65	34	3.22	0.68	34
	3+	3.18	0.93	133	7.92	8.52	36	3.59	0.78	36
	4-	3.46	0.89	63	5.20	4.31	15	3.48	0.72	15
	4	3.49	0.81	55	4.88	4.36	16	3.54	1.02	16
	4+	3.09	1.10	33	3.43	1.13	7	3.50	0.96	7
Calculus AB	2-	2.24	1.28	70	6.40	6.18	48	2.72	0.86	48
	2	2.35	1.29	91	6.34	4.28	56	2.48	1.09	56
	2+	2.60	1.19	92	5.82	3.04	61	2.60	1.05	61
	3-	2.53	1.18	192	6.78	5.28	126	2.60	1.00	126
	3	2.61	1.15	193	6.92	5.23	130	2.64	1.09	130
	3+	2.70	1.18	207	7.22	4.03	135	2.64	1.09	135
	4-	2.85	1.17	202	7.20	4.88	141	2.65	1.03	141
	4	2.94	1.14	233	7.47	4.58	160	2.86	1.04	160
	4+	2.96	1.19	250	7.73	5.82	185	2.84	1.02	185
Biology	2-	2.61	1.04	18	4.07	2.59	14	2.81	0.79	14
	2	2.57	0.95	23	5.23	2.83	13	2.28	1.17	13
	2+	2.54	1.18	24	5.07	3.03	15	2.80	0.70	15
	3-	2.44	1.06	45	3.93	3.23	30	2.77	0.95	30
	3	2.51	1.18	63	3.54	2.29	41	2.83	0.87	41
	3+	2.72	1.17	67	4.34	3.14	47	2.76	1.11	47
	4-	2.95	1.15	65	4.36	3.00	45	3.09	0.74	45
	4	2.76	1.09	72	5.23	3.50	44	3.07	0.99	44
	4+	3.24	1.03	79	4.02	2.69	45	3.39	0.83	45

and the Biology Examination, respectively. Once again the number of students in a number of the score groups is quite small. While ANOVAs were conducted by test year to determine significant differences between the means, it was decided it would be better to combine the information across years in order to obtain results based on larger samples.

Table 11 presents the same information as Tables 8, 9, and 10, but collapsed across test administration dates. This was possible because the method used was similar to equipercentile equating was used in determining equivalent cut scores on the composite score for each of the new AP score groups. For all three of the AP Examinations, the ANOVA yielded statistically significant differences for the average grades in the sequent course (English Language and Composition Examination, $F = 3.86$, $p < .0002$; Calculus AB Examination, $F = 5.85$, $p <. 0001$; Biology Examination, $F = 2.47$, $p < .0125$.) For the English Language and Composition Examination, the 4+, 4, 4-, 3+, and 3 AP score groups earned significantly higher average grades than did the 2- AP score group. As was the case with the middle-third method, the average grades for all AP score groups, except the 2- group, were in the B range. For the Calculus AB Examination, the 4+ AP group differed significantly from the 3, 3-, 2, and 2- AP groups. The 4 AP group also earned a significantly higher average grade in the sequent course that did the 3, 2, and 2- AP groups. For the Biology Examination, the 4+ AP group earned a higher average grade in the sequent course than did the AP group of 3. The 4+ AP group also earned a significantly higher average grade in other Biology classes than did the 3 and 2 AP groups.

As was the case with the results for the middle third method, the 3- AP group did not significantly differ from the other 3 AP groups. Neither did the 3+, 3, or 3- AP groups differ significantly from the 2+ or 4- AP groups on any of the academic outcome measures.

TABLE 8

Descriptive Statistics for the English Outcome Measures by Test Year and AP Score Groups Created by the Middle 50 Percent Method

Test	*AP*	*E 306 Grades*			*Other English Hours Taken*			*GPAs in Other English Classes*		
Date	*Score*	*Mean*	*SD*	*Freq.*	*Mean*	*SD*	*Freq.*	*Mean*	*SD*	*Freq.*
1995	2-	3.41	0.71	17	9.60	9.58	5	3.32	0.46	5
	2	3.08	0.88	61	6.60	7.45	15	2.98	1.06	15
	2+	3.19	0.75	16	6.00	6.00	4	3.30	0.48	4
	3-	3.12	0.93	17	3.00	0.00	4	3.25	0.96	4
	3	3.32	0.82	37	5.63	7.42	8	3.23	0.44	8
	3+	3.30	0.67	10	4.50	2.12	2	4.00	0.00	2
	4-	3.67	0.58	3	3.00		1	3.00		1
	4	3.64	0.67	11	3.00	0.00	5	3.80	0.45	5
	4+	4.00	0.00	2	3.00		1	4.00		1
1996	2-	2.54	0.88	13	4.50	1.73	4	3.50	0.41	4
	2	3.14	0.76	58	7.69	7.67	16	3.40	0.75	16
	2+	3.18	0.76	39	5.00	4.24	9	3.19	1.27	9
	3-	3.13	0.70	48	6.67	8.19	9	3.15	0.33	9
	3	3.07	1.01	74	5.17	3.22	18	3.27	1.00	18
	3+	3.32	0.91	31	13.50	12.00	8	3.74	0.43	8
	4-	3.26	1.05	19	4.20	2.68	5	3.80	0.45	5
	4	3.53	0.75	34	6.00	5.29	10	3.36	1.25	10
	4+	3.63	0.52	8	3.00		1	4.00		1
1997	2-	2.61	1.09	18	6.00	2.45	4	3.46	0.42	4
	2	2.64	1.04	39	4.29	2.36	7	3.31	0.41	7
	2+	3.10	0.88	30	3.00	0.00	6	3.00	0.89	6
	3-	2.90	0.76	30	4.13	3.18	8	2.78	0.98	8
	3	3.18	0.90	73	5.44	6.50	16	3.52	0.49	16
	3+	3.27	0.87	30	6.46	7.22	13	3.70	0.54	13
	4-	3.54	0.66	13	8.40	6.15	5	3.44	0.84	5
	4	2.80	1.44	25	3.00	0.00	5	3.40	0.89	5
	4+	3.20	0.84	5	4.50	2.12	2	2.75	1.77	2
1998	2-	2.53	0.80	17	3.00	0.00	2	3.00	0.00	2
	2	3.14	0.68	36	3.00	0.00	7	3.14	0.69	7
	2+	2.75	1.14	12	3.00	0.00	3	3.00	1.00	3
	3-	3.40	0.68	20	3.75	1.50	4	3.50	1.00	4
	3	3.03	1.10	38	3.43	1.33	7	2.71	0.95	7
	3+	2.95	0.89	20	7.00	6.93	3	3.60	0.69	3
	4-	3.57	0.53	7			0			0
	4	3.62	0.65	13	3.00	0.00	2	3.50	0.71	2
	4+	3.17	0.75	6			0			0
1999	2-	2.00	1.41	4	3.00		1	3.00		1
	2	3.00	0.89	6			0			0
	2+			0			0			0
	3-	3.00	1.00	3	3.00		1	3.00		1
	3	3.00	0.58	7	3.00	0.00	2	4.00	0.00	2
	3+	3.25	0.96	4			0			0
	4-	4.00	0.00	2			0			0
	4	4.00		1			0			0
	4+	3.50	0.71	2	3.00		1	4.00		1

TABLE 9

Descriptive Statistics for the Mathematics Outcome Measures by Test Year and AP Score Groups Created by the Middle 50 Percent Method

Test Date	*AP Score*	*M408C Grades*			*Other Math Hours Taken*			*GPAs in Other Math Classes*		
		Mean	*SD*	*Freq.*	*Mean*	*SD*	*Freq.*	*Mean*	*SD*	*Freq.*
1995	2-			0			0			0
	2	2.00		1	3.00	0.00	2	0.50	0.71	2
	2+			0			0			0
	3-	3.00		1	11.00		1	3.64		1
	3	1.67	1.53	3	4.00		1	0.00		1
	3+			0			0			0
	4-	3.00		1	7.00		1	3.43		1
	4	2.56	1.33	9	5.00	1.87	5	1.97	1.52	5
	4+	3.33	1.15	3	25.00	26.87	2	2.72	1.11	2
1996	2-	2.00	1.12	9	10.40	11.93	10	3.17	0.42	10
	2	2.32	1.25	19	7.20	3.86	15	2.70	0.79	15
	2+	2.81	0.83	16	7.18	4.26	11	2.62	0.72	11
	3-	2.42	1.44	24	8.50	6.39	16	2.77	0.65	16
	3	2.46	1.03	61	7.48	5.92	42	2.79	1.11	42
	3+	2.38	1.26	32	6.88	3.53	26	2.59	1.01	26
	4-	2.53	1.24	34	9.50	7.18	22	2.43	1.08	22
	4	2.70	1.26	92	9.03	6.53	65	2.80	1.05	65
	4+	3.03	1.17	38	9.11	7.43	28	3.02	1.03	28
1997	2-	2.13	1.50	16	4.67	1.87	9	3.00	0.71	9
	2	2.39	1.28	38	7.83	6.12	23	2.30	0.99	23
	2+	2.67	1.14	18	6.38	2.53	16	2.54	1.33	16
	3-	2.73	0.97	49	8.42	7.63	33	2.67	1.05	33
	3	2.78	1.05	90	8.33	6.01	63	2.78	0.95	63
	3+	2.61	1.06	44	8.18	4.06	33	2.59	1.08	33
	4-	3.00	1.02	30	6.26	3.53	23	3.06	0.94	23
	4	2.83	1.15	104	7.18	3.14	76	2.81	1.08	76
	4+	2.80	1.32	45	9.79	7.28	39	2.95	0.95	39
1998	2-	2.17	1.38	18	5.30	1.89	10	1.69	0.79	10
	2	2.26	1.31	34	6.37	2.19	19	2.52	0.96	19
	2+	2.00	1.60	12	6.14	3.76	7	2.25	1.23	7
	3-	2.27	1.23	30	7.13	4.58	23	2.79	0.74	23
	3	2.47	1.21	70	5.72	2.94	40	2.04	1.16	40
	3+	2.89	1.17	47	6.28	2.90	32	2.88	1.07	32
	4-	2.83	1.36	40	6.96	4.06	25	2.76	1.02	25
	4	2.90	1.12	82	7.18	4.05	60	2.59	1.10	60
	4+	3.04	1.03	52	6.74	3.13	38	2.91	0.89	38
1999	2-	2.80	1.01	15	4.30	1.49	10	2.93	0.86	10
	2	2.51	1.19	39	4.18	1.10	22	2.73	1.13	22
	2+	2.67	1.37	18	4.00	1.10	11	3.22	0.60	11
	3-	2.58	1.32	33	4.36	1.62	22	2.33	1.20	22
	3	2.70	1.27	80	5.55	2.98	47	2.70	1.01	47
	3+	2.79	1.29	28	5.00	1.48	12	2.51	1.26	12
	4-	3.04	1.29	27	5.11	1.88	18	2.82	0.98	18
	4	3.22	1.00	89	5.48	2.06	56	2.80	1.04	56
	4+	3.26	1.02	39	6.25	2.93	28	2.69	0.94	28

TABLE 10

Descriptive Statistics for the Biology Outcome Measures by Test Year and AP Score Groups Created by the Middle 50 Percent Method

Test	*AP*	*BIO 303 Grades*			*Other Biology Hours Taken*			*GPAs in Other Biology Classes*		
Date	*Score*	*Mean*	*SD*	*Freq.*	*Mean*	*SD*	*Freq.*	*Mean*	*SD*	*Freq.*
1995	2-			0			0			0
	2	3.33	0.58	3	2.00	0.00	2	4.00	0.00	2
	2+	2.33	1.53	3	3.50	2.12	2	3.00	1.41	2
	3-	2.00	1.41	2	2.00		1	3.00		1
	3	3.00	0.71	5	3.00	2.00	4	2.79	0.53	4
	3+	3.00	0.00	3	2.00	0.00	2	4.00	0.00	2
	4-	3.13	1.46	8	2.80	1.79	5	3.80	0.45	5
	4	2.57	1.02	14	2.75	1.50	4	3.25	0.50	4
	4+	3.80	0.45	5	2.33	0.58	3	3.67	0.58	3
1996	2-	1.67	0.58	3	2.00	0.00	2	2.50	0.71	2
	2	2.57	0.53	7	8.75	2.87	4	2.61	0.64	4
	2+	1.50	1.00	4	5.00		1	1.80		1
	3-	2.33	1.29	15	2.50	0.97	10	3.24	0.94	10
	3	2.70	1.22	20	2.79	2.08	14	2.64	1.01	14
	3+	3.06	1.00	18	4.18	2.64	11	2.88	1.13	11
	4-	2.88	1.13	8	3.75	2.87	4	2.99	0.29	4
	4	3.16	0.88	32	4.43	3.61	21	3.26	0.78	21
	4+	3.00	1.05	19	2.80	1.48	10	3.48	0.51	10
1997	2-	2.67	1.15	3	4.00	1.41	2	3.20	1.13	2
	2	2.50	1.09	12	5.11	2.71	9	2.12	1.17	9
	2+	3.50	1.00	4	5.00	4.08	4	2.83	0.97	4
	3-	2.86	0.69	7	6.40	5.73	5	3.00	0.62	5
	3	2.25	1.24	28	4.12	3.26	17	2.82	0.93	17
	3+	2.79	0.89	14	6.67	4.18	9	3.03	0.61	9
	4-	2.30	1.34	10	4.86	3.34	7	2.81	1.00	7
	4	2.97	1.30	29	7.11	4.00	18	3.21	1.00	18
	4+	3.33	0.71	9	6.00	4.24	4	2.78	0.52	4
1998	2-	2.75	1.89	4	5.75	3.86	4	3.03	0.50	4
	2	2.71	0.76	7	3.60	1.34	5	2.39	0.99	5
	2+	2.80	0.84	5	5.33	2.52	3	3.01	0.35	3
	3-	2.50	0.85	10	3.75	2.12	8	2.48	0.78	8
	3	2.81	0.87	21	4.07	1.62	15	2.67	1.17	15
	3+	2.33	1.66	9	4.38	4.34	8	2.45	1.68	8
	4-	3.00	0.93	15	3.77	2.68	13	2.84	0.70	13
	4	2.89	1.12	27	5.21	2.72	19	3.32	0.78	19
	4+	3.00	1.58	13	5.33	3.50	6	3.54	1.12	6
1999	2-	2.67	0.58	3	3.00		1	2.33		1
	2	2.50	0.58	4	4.50	0.71	2	2.60	0.85	2
	2+	2.33	2.08	3	3.00		1	2.33		1
	3-	2.67	0.58	3	2.00		1	2.00		1
	3	2.27	1.33	15	3.80	1.75	10	2.87	0.69	10
	3+	1.80	1.79	5	3.67	1.53	3	1.73	0.64	3
	4-	3.00	1.41	8	3.75	1.50	4	2.93	0.39	4
	4	3.25	0.75	12	3.73	1.68	11	2.56	1.46	11
	4+	3.14	0.69	7	3.40	0.89	5	3.71	0.40	5

TABLE 11

Descriptive Statistics for the Course Outcome Measures Where the AP Score Groups Were Created by the Middle 50 Percent Method

AP Exam	*AP Score*	*Next Course Grades*			*Other Hours Taken in Subject Area*			*GPAs in Other Classes Taken in Subject Area*		
		Mean	*SD*	*Freq.*	*Mean*	*SD*	*Freq.*	*Mean*	*SD*	*Freq.*
English Language and Composition	2-	2.74	0.98	69	6.19	5.74	16	3.34	0.40	16
	2	3.02	0.86	200	6.07	6.44	45	3.20	0.82	45
	2+	3.10	0.85	97	4.36	3.67	22	3.13	0.97	22
	3-	3.11	0.76	118	4.73	5.17	26	3.10	0.77	26
	3	3.14	0.95	229	5.00	4.97	51	3.29	0.79	51
	3+	3.22	0.87	95	8.54	9.01	26	3.72	0.49	26
	4-	3.45	0.82	44	6.00	4.84	11	3.56	0.66	11
	4	3.35	1.04	84	4.36	3.79	22	3.48	0.96	22
	4+	3.43	0.66	23	3.60	1.34	5	3.50	1.12	5
Calculus AB	2-	2.29	1.30	58	6.20	6.49	39	2.69	0.91	39
	2	2.38	1.24	131	6.26	4.08	81	2.50	1.03	81
	2+	2.58	1.24	64	5.96	3.15	45	2.68	1.06	45
	3-	2.54	1.21	137	7.21	5.89	95	2.65	0.97	95
	3	2.61	1.15	304	6.91	4.96	193	2.59	1.09	193
	3+	2.68	1.19	151	6.89	3.48	103	2.67	1.08	103
	4-	2.83	1.24	132	7.03	4.80	89	2.78	1.01	89
	4	2.90	1.15	376	7.24	4.42	262	2.74	1.08	262
	4+	3.03	1.14	177	8.28	6.52	135	2.90	0.95	135
Biology	2-	2.46	1.20	13	4.22	2.91	9	2.87	0.65	9
	2	2.64	0.82	33	5.09	2.89	22	2.48	1.04	22
	2+	2.53	1.31	19	4.64	2.73	11	2.77	0.81	11
	3-	2.49	1.02	37	3.64	3.08	25	2.89	0.85	25
	3	2.53	1.16	89	3.67	2.33	60	2.75	0.94	60
	3+	2.71	1.21	49	4.73	3.55	33	2.78	1.17	33
	4-	2.86	1.21	49	3.85	2.56	33	3.01	0.73	33
	4	2.98	1.06	114	5.10	3.39	73	3.16	0.96	73
	4+	3.15	1.08	53	3.86	2.65	28	3.46	0.70	28

Phase I: Cluster Analyses

The results of the cluster analyses for the seven-cluster solutions are presented first, followed by the results for the eleven-cluster solutions. The results that are reported are for the cluster analyses of the multiple-choice and constructed response section scores. While cluster analyses were performed using the item level data, the results were difficult to interpret. For example, it was not uncommon to find for a given cluster analysis that more than 50 percent of the clusters included individuals from all AP score groups. This made it difficult to differentiate the clusters in terms of the AP score groups. As a consequence, it was decided to not report the item level analyses in this report. Readers interested in these results should contact the authors.

Seven-cluster solutions: English Language and Composition Examination. Table 12 presents the average AP score for each of the seven clusters and shows the distribution of AP scores within each cluster for the national sample who took the examination in 1995. Because it was difficult to label the clusters as high and low groups within an AP score group, the mean AP score was selected to differentiate the clusters. As can be seen in Table 12, all the clusters contain students from two or three different AP score groups. Inspection of the cluster centers revealed that the difference between the clusters that contain students from the same AP score group were a function of their performance on the two section scores. For example, students in the 1.1 cluster performed better on the multiple-choice section (cluster center = -1.26) than they did on the constructed response section (cluster center = -1.86), while the opposite was found for students in the 1.8 cluster. Students in the 1.8 cluster had a higher cluster center on the constructed response section (-.50) than they did on the multiple-choice section (-1.29). Similar results were found when students in a given AP score group were divided into two clusters. When students within a given AP score group were distributed across three clusters, it was found that the cluster centers represented better performance on one of the two section scores or the same level of performance on both sections.

TABLE 12

English 1995: Seven National Clusters Against AP Scores

Mean AP Score Within Cluster		*AP Score* 1	2	3	4	5	*Total*
1.1	Number	4,013	379				4,392
	Column %	75.5%	2.0%				8.6%
1.8	Number	1,281	5,687				6,968
	Column %	24.1%	30.0%				13.6%
2.1	Number	19	7,264	553			7,836
	Column %	0.4%	38.4%	3.7%			15.3%
2.4	Number		5,596	3,833		1	9,430
	Column %		29.6%	25.4%		0.0%	18.4%
3.3	Number			6,525	2,290		8,815
	Column %			43.3%	27.0%		17.2%
3.4	Number			4,150	3,009	19	7,178
	Column %			27.6%	35.5%	0.5%	14.0%
4.5	Number				3,172	3,512	6,684
	Column %				37.4%	99.4%	13.0%
Total	Number	5,313	18,926	15,061	8,471	3,532	51,303

TABLE 13

English 1996: Seven National Clusters Against AP Scores

Mean AP Score Within Cluster		*AP Score* 1	2	3	4	5	*Total*
1.2	Number	3,892	735				4,627
	Column %	93.4%	4.0%				7.8%
2.0	Number	50	8,231	90			8,371
	Column %	1.2%	45.0%	0.4%			14.2%
2.0	Number	225	5,201	476			5,902
	Column %	5.4%	28.4%	2.4%			10.0%
2.6	Number		4,141	5,813			9,954
	Column %		22.6%	29.0%			16.9%
3.3	Number			8,300	2,901		11,201
	Column %			41.3%	23.7%		19.0%
3.5	Number			5,400	5,298	40	10,738
	Column %			26.9%	43.3%	1.0%	18.2%
4.5	Number				4,045	4,143	8,188
	Column %				33.0%	99.0%	13.9%
Total	Number	4,167	18,308	20,079	12,244	4,183	58,981

TABLE 14

English 1997: Seven National Clusters Against AP Scores

Mean AP Score Within Cluster		*AP Score* 1	2	3	4	5	*Total*
1.5	Number	3,925	3,564				7,489
	Column %	98.8%	18.4%				11.2%
2.1	Number	19	8,795	1,348			10,162
	Column %	0.5%	45.4%	5.6%			15.2%
2.2	Number	27	5,971	1,884			7,882
	Column %	0.7%	30.8%	7.8%			11.8%
2.9	Number		1,042	9,690			10,732
	Column %		5.4%	40.3%			16.0%
3.4	Number			6,875	3,956	22	10,853
	Column %			28.6%	31.6%	0.3%	16.2%
3.6	Number			4,261	5,106	319	9,686
	Column %			17.7%	40.7%	4.5%	14.5%
4.7	Number				3,474	6,683	10,157
	Column %				27.7%	95.1%	15.2%
Total	Number	3,971	19,372	24,058	12,536	7,024	66,961

TABLE 15

English 1998: Seven National Clusters Against AP Scores

Mean AP Score Within Cluster		*AP Score* 1	2	3	4	5	*Total*
1.2	Number	4,692	1,200				5,892
	Column %	94.6%	5.2%				7.4%
2.0	Number	151	10,848				10,999
	Column %	3.0%	47.4%				13.8%
2.1	Number	119	5,799	655			6,573
	Column %	2.4%	25.3%	2.4%			8.3%
2.6	Number		4,869	7,949			12,818
	Column %		21.3%	29.5%			16.1%
3.2	Number		173	11,310	2,849		14,332
	Column %		0.8%	42.0%	16.6%		18.0%
3.5	Number			7,021	7,202	156	14,379
	Column %			26.1%	42.0%	2.1%	18.1%
4.5	Number				7,101	7,352	14,453
	Column %				41.4%	97.9%	18.2%
Total	Number	4,962	22,889	26,935	17,152	7,508	79,446

Tables 13–16 present the results of cluster analysis of the national samples for the 1996–1999 test administrations, respectively. The distribution of AP Scores within clusters for these administrations appears very similar to those found for the 1995 administration. The pattern for the cluster centers found for 1995 was also replicated in the 1996–1999 test administrations.

Table 17 presents the descriptive statistics of the English outcome measures for the 1995 University of Texas sample that was classified according to the cluster centers obtained from the analysis of the 1995 national data for the English Language and Composition Examination. The outcome measures include the grade in the next English course (E 316K), other English hours taken, and the GPA in the other English courses. Clusters are labeled according to the AP means for the national clusters not the University of Texas sample. An ANOVA yielded a significant difference among the clusters for the grades in E 316K ($F = 8.19$, $p < .001$). Students in the 4.5 cluster earned a significantly higher average grade in E 316K than did students in the all of the clusters with an AP mean less than or equal to 2.4. The students in the 3.4 cluster also had a statistically significant higher average grade in E 316K than did students in the 1.1 and 1.8 clusters. The clusters also differed statistically for the number of additional English hours taken ($F = 9.29$). Students in the 4.5 cluster took significantly more hours than did students in all of the other clusters. The clusters that differ significantly are noted at the bottom of the table.

The descriptive statistics for the English outcome

TABLE 16

English 1999: Seven National Clusters Against AP Scores

Mean AP Score Within Cluster		*AP Score* 1	2	3	4	5	*Total*
1.1	Number	4,514	5,294				9,808
	Column %	99.8%	16.6%				10.1%
2.1	Number		15,065	1,271			16,336
	Column %		47.1%	3.8%			16.9%
2.2	Number	7	9,622	2,275			11,904
	Column %	0.2%	30.1%	6.7%			12.3%
2.8	Number		1,990	15,790			17,780
	Column %		6.2%	46.6%			18.4%
3.4	Number			9,276	5,756	44	15,076
	Column %			27.4%	32.6%	0.5%	15.6%
3.7	Number			5,249	7,507	653	13,409
	Column %			15.5%	42.5%	7.5%	13.9%
4.6	Number				4,391	7,999	12,390
	Column %				24.9%	92.0%	12.8%
Total	Number	4,521	31,971	33,861	17,654	8,696	96,703

TABLE 17

English 1995: Descriptive Statistics of the English Outcome Measures

	Mean AP Score Within Cluster	*Frequency*	*Mean*	*Standard Deviation*
E 316K Grades	1.1	45	3.22	0.90
	1.8	80	3.25	0.74
	2.1	54	3.33	0.67
	2.4	73	3.33	0.75
	3.3	67	3.61	0.49
	3.4	45	3.60	0.58
	4.5	80	3.83	0.44
Other English Hours Taken	1.1	15	6.20	6.26
	1.8	33	5.45	6.16
	2.1	31	9.90	8.98
	2.4	53	6.00	6.36
	3.3	54	9.11	8.97
	3.4	41	8.27	9.31
	4.5	59	17.14	13.35
GPA in Other English Classes	1.1	15	3.15	0.82
	1.8	33	3.41	0.86
	2.1	31	3.41	0.82
	2.4	53	3.44	0.75
	3.3	54	3.39	0.80
	3.4	41	3.59	0.91
	4.5	59	3.71	0.42

E 316K Grades: 4.5 > 1.1 – 2.4; 3.3 > 1.1, 1.8 Other Hours: 4.5 > 1.1 – 3.4

measures for the 1996 test administration of the English Language and Composition Examination for the University of Texas at Austin sample are shown in Table 18. The ANOVA yielded statistically significant differ-

TABLE 18

English 1996: Descriptive Statistics of the English Outcome Measures

	Mean AP Score Within Cluster	*Frequency*	*Mean*	*Standard Deviation*
E 316K Grades	1.2	73	3.12	0.91
	2.0(a)	133	3.23	0.70
	2.0(b)	83	3.30	0.74
	2.6	107	3.24	0.91
	3.3	96	3.45	0.77
	3.5	89	3.61	0.58
	4.5	97	3.77	0.59
Other English Hours Taken	1.2	26	6.46	7.18
	2.0(a)	38	6.32	6.03
	2.0(b)	27	6.56	7.94
	2.6	61	7.72	7.62
	3.3	68	7.87	8.19
	3.5	74	9.93	7.71
	4.5	92	14.54	10.66
GPAs in Other English Classes	1.2	26	3.50	0.62
	2.0(a)	38	3.29	0.78
	2.0(b)	27	3.22	1.18
	2.6	61	3.26	0.74
	3.3	68	3.50	0.75
	3.5	74	3.39	0.98
	4.5	92	3.65	0.64

E 316K Grades: 4.5 > 1.2 – 3.3; 3.5 > 1.2, 2.0(a), 2.6 Other Hours: 4.5 > 1.2 – 3.5 Other GPAs: 4.5 > 2.6

ences for the grades in E 316K ($F = 9.00$, $p < .001$), other English hours taken ($F = 8.41$, $p < .001$), and the GPA in other English classes ($F = 2.33$, $p < .04$). The clusters that differed significantly from one another are listed at the bottom of the table. It should be noted the clusters labeled with a 3 did not differ significantly from one another.

Table 19 presents the same information as Table 18 but for the 1997 test administration. Significant differences were found for the grades in E 316K ($F = 12.43$, $p < .001$), other English hours taken ($F = 8.17$, $p < .001$), and GPA in other English classes ($F = 3.81$, $p < .001$). Inspection of the significant differences that are listed at the bottom of the table reveals the clusters labeled 3 do not differ significantly from one another.

Descriptive statistics of the English outcome measures for the 1998 test date for the University of Texas at Austin sample clustered according to the cluster centers derived from the 1998 national sample are presented in Table 20. The ANOVA yielded significant differences for the grades in E 316K ($F = 13.01$; $p < .001$), other English hours taken ($F = 11.06$; $p < .001$), and GPA in other English classes ($F = 6.53$, $p < .001$). The direction of the significant mean differences appears at the bottom of the table.

Table 21 presents the same information as Table 20 but for the 1999 test administration. For this year only the E 316K grades differed significantly ($F = 3.54$, $p < .004$). Cluster 4.6 earned a higher average grade in E 316K than did clusters 1.1 and 2.1.

Seven-cluster solutions: Calculus AB Examination. Tables 22–26 present AP score distribution for the clusters obtained from the cluster analysis of the national samples for the 1995–1999 test administrations, respectively. The results are very similar to the patterns observed in the analysis of the English Language and Composition Examination. The multiple clusters found for a given AP score group differ from one another in terms of performance on the sections. The center clusters revealed that students within a cluster performed better on one of the two sections of the examination or about equal.

The descriptive statistics of the outcome measures in calculus for the University of Texas at Austin sample who took the examination in 1995 are presented in Table 27. No ANOVAs were conducted because of the small number of students in each cluster.

Table 28 presents the same information as Table 27 but for the 1996 test administration date. Significant differences were found among the clusters for M 408D grades ($F = 6.69$, $p < .001$), other math hours taken ($F = 2.64$, $p < .03$), and GPA in other classes ($F = 5.51$, $p < .001$). How the various clusters differed is listed at the bottom of the table.

TABLE 19

English 1997: Descriptive Statistics of the English Outcome Measures

	Mean AP Score Within Cluster	*Frequency*	*Mean*	*Standard Deviation*
E 316K Grades	1.5	69	2.88	0.99
	2.1	88	3.23	0.64
	2.2	83	3.01	1.01
	2.9	101	3.47	0.56
	3.4	60	3.28	0.94
	3.6	76	3.46	0.82
	4.7	107	3.75	0.52
Other English Hours Taken	1.5	16	4.50	2.19
	2.1	31	4.45	3.70
	2.2	55	5.78	5.53
	2.9	61	5.56	4.83
	3.4	64	5.63	4.71
	3.6	45	8.73	7.40
	4.7	98	10.04	7.21
GPAs in Other English Classes	1.5	16	3.31	0.55
	2.1	31	3.23	0.85
	2.2	55	3.16	0.90
	2.9	61	3.41	0.87
	3.4	64	3.61	0.55
	3.6	45	3.55	0.69
	4.7	98	3.65	0.65

E 316K Grades: 4.7 > 2.1, 3.4, 2.2; 2.9, 3.6 > 2.2; 2.9, 3.6, 4.7 > 1.5 Other Hours: 4.7 > 1.5 – 3.4; 3.6 > 2.1 Other GPAs: 4.7, 3.4 > 2.2

TABLE 20

English 1998: Descriptive Statistics of the English Outcome Measures

	Mean AP Score Within Cluster	*Frequency*	*Mean*	*Standard Deviation*
E 316K Grades	1.2	54	3.13	0.78
	2.0	48	3.21	0.82
	2.1	36	3.08	0.60
	2.6	107	3.25	0.89
	3.2	73	3.51	0.58
	3.5	121	3.62	0.61
	4.5	115	3.82	0.47
Other English Hours Taken	1.2	13	3.00	0.00
	2.0	24	4.25	3.05
	2.1	11	3.27	0.90
	2.6	42	3.79	1.88
	3.2	49	5.14	3.92
	3.5	71	6.30	4.12
	4.5	79	8.20	4.40
GPAs in Other English Classes	1.2	13	3.08	0.64
	2.0	24	3.24	0.94
	2.1	11	3.27	0.90
	2.6	42	3.25	0.78
	3.2	49	3.63	0.46
	3.5	71	3.65	0.59
	4.5	79	3.78	0.43

E 316K Grades: 4.5 > 1.2 – 3.2; 3.5 > 1.2 – 2.6; 3.2 > 1.2, 2.1 Other Hours: 4.5 > 1.2 – 3.5; 3.5 > 1.2, 2.6
Other GPAs: 4.5 > 1.2, 2.0, 2.6; 3.5 > 1.2, 2.6

TABLE 21

English 1999: Descriptive Statistics of the English Outcome Measures

	Mean AP Score Within Cluster	*Frequency*	*Mean*	*Standard Deviation*
E 316K Grades	1.1	5	2.60	1.52
	2.1	9	2.89	0.93
	2.2	10	3.10	0.74
	2.8	11	3.36	0.67
	3.4	10	3.60	0.52
	3.7	19	3.53	0.51
	4.6	12	3.92	0.29
Other English Hours Taken	1.1	2	3.00	0.00
	2.1	0		
	2.2	4	3.00	0.00
	2.8	9	3.00	0.00
	3.4	14	4.29	3.27
	3.7	14	4.71	3.47
	4.6	10	7.80	4.73
GPAs in Other English Classes	1.1	2	3.50	0.71
	2.1	0		
	2.2	4	3.75	0.50
	2.8	9	3.22	1.30
	3.4	14	3.36	0.72
	3.7	14	3.30	0.90
	4.6	10	3.81	0.26

TABLE 22

Calculus 1995: Seven National Clusters Against AP Scores

Mean AP Score Within Cluster		*AP Score* 1	2	3	4	5	*Total*
1.0	Number	15,060					15,060
	Column %	64.7%					14.6%
1.6	Number	8,166	9,993	2			18,161
	Column %	35.1%	52.7%	0.0%			17.7%
2.5	Number	50	8,074	7,118			15,242
	Column %	0.2%	42.6%	24.7%			14.8%
3.0	Number		883	15,349	426		16,558
	Column %		4.7%	52.8%	2.2%		16.1%
3.5	Number			6,507	7,520	50	14,077
	Column %			22.5%	38.4%	0.4%	13.7%
4.2	Number				11,644	2,440	14,084
	Column %				59.4%	20.1%	13.7%
5.0	Number					9,638	9,638
	Column %					79.5%	9.4%
Total	Number	23,276	18,950	28,876	19,590	12,128	102,820

TABLE 23

Calculus 1996: Seven National Clusters Against AP Scores

Mean AP Score Within Cluster		*AP Score* 1	2	3	4	5	*Total*
1.0	Number	17,296	167				17,463
	Column %	85.2%	0.8%				16.6%
1.9	Number	2,996	15,609	674			19,279
	Column %	14.8%	74.0%	2.4%			18.4%
2.6	Number	2	5,294	9,184			14,480
	Column %	0.0%	25.1%	32.4%			13.8%
3.2	Number		12	14,577	2,620		17,209
	Column %		0.1%	51.4%	12.0%		16.4%
3.7	Number			3,928	9,726	139	13,793
	Column %			13.8%	44.5%	1.0%	13.1%
4.3	Number				9,525	4,434	13,959
	Column %				43.6%	33.0%	13.3%
5.0	Number					8,857	8,857
	Column %					65.9%	8.4%
Total	Number	20,294	21,082	28,363	21,871	13,430	105,040

TABLE 24

Calculus 1997: Seven National Clusters Against AP Scores

Mean AP Score Within Cluster		*AP Score* 1	2	3	4	5	*Total*
1.0	Number	16,214					16,214
	Column %	71.1%					14.6%
1.7	Number	6,472	12,045				18,517
	Column %	28.4%	54.4%				16.7%
2.4	Number	109	8,576	5,043			13,728
	Column %	0.5%	38.8%	16.6%			12.3%
2.9	Number		1,501	19,406	14		20,921
	Column %		6.8%	63.9%	0.1%		18.8%
3.6	Number			5,899	8,219	8	14,126
	Column %			19.4%	36.2%	0.1%	12.7%
4.1	Number				14,203	1,521	15,724
	Column %				62.6%	11.5%	14.1%
5.0	Number				250	11,722	11,972
	Column %				1.1%	88.5%	10.8%
Total	Number	22,795	22,122	30,348	22,686	13,251	111,202

TABLE 25

Calculus 1998: Seven National Clusters Against AP Scores

Mean AP Score Within Cluster		*AP Score* 1	2	3	4	5	*Total*
1.0	Number	15,476	28				15,504
	Column %	82.0%	0.1%				13.3%
1.8	Number	3,400	14,445				17,845
	Column %	18.0%	69.7%				15.3%
2.6	Number		5,896	8,880			14,776
	Column %		28.4%	28.4%			12.7%
3.1	Number		363	15,768	2,009		18,140
	Column %		1.8%	50.4%	7.4%		15.6%
3.6	Number			6,639	8,992		15,631
	Column %			21.2%	33.2%		13.4%
4.2	Number				16,052	4,448	20,500
	Column %				59.2%	24.0%	17.6%
5.0	Number				50	14,074	14,124
	Column %				0.2%	76.0%	12.1%
Total	Number	18,876	20,732	31,287	27,103	18,522	116,520

All of the differences reflect the higher average AP scores obtained by the 5.0 and 4.3 clusters relative to the other clusters.

Significant differences were found among the clusters in terms of average M 408D grades ($F = 10.77$, $p < .001$) and GPA in other mathematics courses taken ($F = 8.12$, $p < .001$) for the 1997 test administration for the University of Texas at Austin sample. Table 29 presents the descriptive statistics of the outcome measures and lists the clusters that differed significantly. Again the statistically significant differences typically involved the extreme clusters.

Like the results for 1997, the 1998 and 1999 test administration analyses yielded statistically significant cluster differences for the average M 408D grades (1998, $F = 9.39$, $p < .001$; 1999, $F = 18.28$, $p < .001$) and the GPA in other mathematics courses (1998, $F = 4.25$, $p < .001$; 1999, $F = 5.40$, $p < .001$). Tables 30 and 31 show the descriptive statistics of the calculus outcome measures.

Seven-cluster solutions: Biology Examination. Tables 32–36 show the distribution of AP scores for each of the clusters obtained from the cluster analysis of each of the 1995–1999 test administrations of the AP Biology Examination for the national samples, respectively. The results are similar to those found for the other two AP Examinations.

The descriptive statistics for the Biology outcome measures for each of the 1995–1999 test administration years for the University of Texas samples are presented in Tables 37–41, respectively. No statistical significance tests were conducted on the 1995 sample

TABLE 26

Calculus 1999: Seven National Clusters Against AP Scores

Mean AP Score Within Cluster		*AP Score* 1	2	3	4	5	*Total*
1.0	Number	20,408					20,408
	Column %	94.3%					16.2%
2.0	Number	1,226	21,789	1,186			24,201
	Column %	5.7%	90.8%	3.8%			16.2%
2.9	Number		2,194	16,742			18,936
	Column %		9.1%	53.4%			15.1%
3.3	Number		17	11,460	5,319		16,796
	Column %		0.1%	36.6%	18.6%		13.4%
3.9	Number			1,951	14,356	60	16,367
	Column %			6.2%	50.3%	0.3%	13.0%
4.5	Number				8,860	9,523	18,383
	Column %				31.0%	47.0%	14.6%
5.0	Number					10,688	10,688
	Column %					52.7%	8.5%
Total	Number	21,634	24,000	31,339	28,535	20,271	125,779

TABLE 27

Calculus 1995: Descriptive Statistics of the Mathematics Outcome Measures

	Mean AP Score Within Cluster	*Frequency*	*Mean*	*Standard Deviation*
M 408D Grades	1.0	1	2.00	
	1.6	4	2.00	1.41
	2.5	6	3.00	1.26
	3.0	8	2.63	1.19
	3.5	9	2.89	1.05
	4.2	7	3.43	0.79
	5.0	5	4.00	0.00
Other Math Hours Taken	1.0	6	5.00	2.37
	1.6	6	5.50	3.02
	2.5	11	11.18	11.64
	3.0	10	4.30	1.49
	3.5	10	5.80	2.44
	4.2	9	12.00	15.55
	5.0	5	13.40	8.14
GPAs in Other Math Classes	1.0	6	2.00	1.90
	1.6	6	2.42	1.61
	2.5	11	3.12	1.12
	3.0	10	2.91	1.30
	3.5	10	3.21	0.93
	4.2	9	3.45	0.73
	5.0	5	3.81	0.31

TABLE 28

Calculus 1996: Descriptive Statistics of the Mathematics Outcome Measures

	Mean AP Score Within Cluster	*Frequency*	*Mean*	*Standard Deviation*
M 408D Grades	1.0	32	2.25	0.98
	1.9	62	2.55	1.14
	2.6	44	2.32	1.27
	3.2	85	2.61	1.23
	3.7	95	2.51	1.32
	4.3	70	3.20	0.97
	5.0	78	3.21	1.23
Other Math Hours Taken	1.0	95	7.73	5.27
	1.9	97	6.67	4.75
	2.6	72	6.21	4.41
	3.2	102	7.05	5.50
	3.7	97	7.91	6.49
	4.3	79	7.09	5.32
	5.0	80	9.45	8.52
GPAs in Other Math Classes	1.0	95	2.69	0.94
	1.9	97	3.00	0.88
	2.6	72	3.03	1.07
	3.2	102	3.01	0.94
	3.7	97	2.91	1.16
	4.3	79	3.50	0.65
	5.0	80	3.16	1.09

M 408D Grades: 4.3, 5.0 > 1.0 – 3.7 Other Hours: 5.0 > 1.9, 2.6 Other GPAs: 5.0 > 1.0; 4.3 > 1.0 – 3.7

TABLE 29

Calculus 1997: Descriptive Statistics of the Mathematics Outcome Measures

	Mean AP Score Within Cluster	*Frequency*	*Mean*	*Standard Deviation*
M 408D Grades	1.0	33	2.00	1.46
	1.7	73	2.38	1.28
	2.4	87	2.78	0.97
	2.9	91	2.66	1.09
	3.6	131	2.95	1.19
	4.1	76	2.83	1.08
	5.0	86	3.51	0.97
Other Math Hours Taken	1.0	77	6.05	2.79
	1.7	109	6.89	3.71
	2.4	105	7.24	6.03
	2.9	108	6.81	4.30
	3.6	152	6.27	3.95
	4.1	92	7.30	5.36
	5.0	89	7.63	5.41
GPAs in Other Math Classes	1.0	77	2.39	1.13
	1.7	109	2.76	1.01
	2.4	105	2.71	1.10
	2.9	108	2.96	1.00
	3.6	152	3.06	0.97
	4.1	92	2.89	1.10
	5.0	89	3.38	0.90

M 408D Grades: 5.0 > 1.0 – 4.1; 2.4, 3.6, 4.1 > 1.0; 3.6 > 1.7 Other GPAs: 5.0 > 1.0 – 2.4, 4.1; 2.9, 3.6, 4.1 > 1.0

TABLE 30

Calculus 1998: Descriptive Statistics of the Mathematics Outcome Measures

	Mean AP Score Within Cluster	*Frequency*	*Mean*	*Standard Deviation*
M 408D Grades	1.0	31	2.52	1.41
	1.8	59	2.07	1.35
	2.6	58	2.47	1.23
	3.1	80	2.65	1.27
	3.6	99	2.89	1.16
	4.2	128	2.92	1.15
	5.0	94	3.38	0.80
Other Math Hours Taken	1.0	95	6.03	2.84
	1.8	80	6.15	2.66
	2.6	73	6.29	3.08
	3.1	83	5.34	2.46
	3.6	111	6.46	3.77
	4.2	126	6.37	3.30
	5.0	83	7.05	4.49
GPAs in Other Math Classes	1.0	95	2.70	1.02
	1.8	80	2.86	1.02
	2.6	73	2.90	0.87
	3.1	83	2.78	1.16
	3.6	111	3.00	1.00
	4.2	126	2.93	1.11
	5.0	83	3.40	0.80

M 408D Grades: 5.0 > 1.0 – 3.1; 4.2 > 1.8 Other GPAs: 5.0 > 1.0 – 3.1, 4.2

TABLE 31

Calculus 1999: Descriptive Statistics of the Mathematics Outcome Measures

	Mean AP Score Within Cluster	*Frequency*	*Mean*	*Standard Deviation*
M 408D Grades	1.0	51	1.82	1.34
	2.0	81	2.46	1.33
	2.9	86	2.70	1.21
	3.2	64	2.97	1.17
	3.9	92	3.27	1.03
	4.5	102	3.16	1.11
	5.0	70	3.70	0.75
Other Math Hours Taken	1.0	106	4.75	2.31
	2.0	114	5.11	2.05
	2.9	87	5.18	2.65
	3.2	65	4.88	1.75
	3.9	98	5.05	1.96
	4.5	106	4.92	2.09
	5.0	61	5.28	1.83
GPAs in Other Math Classes	1.0	106	2.82	1.08
	2.0	114	3.00	0.94
	2.9	87	2.74	1.12
	3.2	65	3.09	0.95
	3.9	98	3.01	1.03
	4.5	106	3.20	0.88
	5.0	61	3.54	0.86

M 408D Grades: 5.0 > 1.0 – 3.2, 4.5; 4.5 > 2.0; 3.9 > 2.0 – 2.9; 2.0 – 5.0 > 1.0; 2.0 > 1.0 Other GPAs: 5.0 > 1.0 – 2.9, 3.9; 4.5 > 2.9

because of the small number of students in each of the clusters. Significant differences were found for the BIO 303 grades for the remaining four test administrations (1996, F = 15.02, p < .001; 1997, F = 24.09, p < .001; 1998, F = 15.07, p < .001; 1999, F = 15.50, p < .001). For three of the test administration dates, significant differences were found for the GPA in other biology courses (1996, F = 6.33, p < .001; 1997, F = 3.20, p < .005; 1999, F = 2.98, p < .009). As was the case with the other AP Examinations, the clusters that differ significantly from one another are listed at the bottom of the table for each year.

Eleven-cluster solutions. The same cluster analyses that were described in the previous three sections were

TABLE 32

Biology 1995: Seven National Clusters Against AP Scores

Mean AP Score Within Cluster		*AP Score* 1	2	3	4	5	*Total*
1.0	Number	6,423		1			6,426
	Column %	78.1%		0.0%			10.5%
1.8	Number	1,799	9,727				11,524
	Column %	21.9%	69.8%				18.9%
2.6	Number		3,554	6,001			9,555
	Column %		25.5%	39.6%			15.6%
3.1	Number		652	7,266	912		8,830
	Column %		4.7%	48.0%	6.9%		14.4%
3.8	Number			1,870	6,449	239	8,558
	Column %			12.4%	48.6%	2.3%	14.0%
4.4	Number				5,908	3,542	9,450
	Column %				44.5%	33.6%	15.5%
5.0	Number					6,769	6,769
	Column %					64.2%	11.1%
Total	Number	8,222	13,933	15,138	13,269	10,550	61,112

TABLE 33

Biology 1996: Seven National Clusters Against AP Scores

Mean AP Score Within Cluster		*AP Score* 1	2	3	4	5	*Total*
1.0	Number	7,575					7,575
	Column %	81.1%					11.5%
1.9	Number	1,764	10,529				12,293
	Column %	18.9%	73.4%				18.7%
2.8	Number		2,183	6,478	51		8,712
	Column %		15.2%	42.2%	0.4%		13.3%
2.9	Number		1,628	8,311	417		10,356
	Column %		11.4%	54.2%	2.9%		15.8%
4.1	Number			497	7,429	1,528	9,454
	Column %			3.2%	51.9%	12.4%	14.4%
4.2	Number			62	6,404	1,981	8,447
	Column %			0.4%	44.8%	16.1%	12.9%
5.0	Number					8,789	8,789
	Column %					71.5%	13.4%
Total	Number	9,339	14,340	15,348	14,301	12,298	65,626

TABLE 34

Biology 1997: Seven National Clusters Against AP Scores

Mean AP Score Within Cluster		*AP Score* 1	2	3	4	5	*Total*
1.1	Number	6,718	452				7,170
	Column %	83.2%	3.0%				10.2%
1.9	Number	1,358	11,233				12,591
	Column %	16.8%	75.7%				17.8%
2.8	Number		1,932	6,834	114		8,880
	Column %		13.0%	38.3%	0.7%		12.6%
3.0	Number		1,213	10,660	764		12,637
	Column %		8.2%	59.8%	4.9%		17.9%
4.2	Number			274	7,869	2,111	10,254
	Column %			1.5%	50.0%	15.0%	14.5%
4.2	Number			59	7,005	2,363	9,427
	Column %			0.3%	44.5%	16.8%	13.4%
5.0	Number					9,612	9,612
	Column %					68.2%	13.6%
Total	Number	8,076	14,830	17,827	15,752	14,086	70,571

TABLE 35

Biology 1998: Seven National Clusters Against AP Scores

Mean AP Score Within Cluster		*AP Score* 1	2	3	4	5	*Total*
1.0	Number	8,011					8,011
	Column %	65.7%					10.6%
1.7	Number	4,176	9,496				13,672
	Column %	34.3%	55.0%				18.1%
2.4	Number	3	7,719	6,079			13,801
	Column %	0.0%	44.7%	33.9%			18.3%
3.3	Number		56	7,844	3,319	2	11,221
	Column %		0.3%	43.8%	23.9%	0.0%	14.9%
3.6	Number			3,991	6,328	64	10,383
	Column %			22.3%	45.5%	0.5%	13.8%
4.6	Number				4,246	6,675	10,921
	Column %				30.6%	47.0%	14.5%
5.0	Number					7,456	7,456
	Column %					52.5%	9.9%
Total	Number	12,190	17,271	17,914	13,893	14,197	75,465

TABLE 36

Biology 1999: Seven National Clusters Against AP Scores

Mean AP Score Within Cluster		*AP Score* 1	2	3	4	5	*Total*
1.1	Number	10,631	1,334				11,965
	Column %	99.9%	7.5%				14.7%
2.1	Number	7	16,002	1,830			17,839
	Column %	0.1%	90.1%	9.6%			21.9%
3.2	Number		181	8,864	2,398		11,443
	Column %		1.0%	46.3%	13.3%		14.0%
3.6	Number		249	8,453	4,032		12,734
	Column %		1.4%	44.1%	22.4%		15.6%
4.4	Number				7,072	3,838	10,910
	Column %				39.3%	24.0%	13.4%
4.5	Number				4,476	4,491	8,967
	Column %				24.9%	28.1%	11.0%
5.0	Number					7,634	7,634
	Column %					47.8%	9.4%
Total	Number	10,638	17,766	19,147	17,978	15,963	81,492

conducted again for the section scores for each of the AP Examinations except that eleven clusters were specified instead of seven clusters. The same tables were prepared and are presented in the Appendix. Generally, the eleven cluster solutions for each examination yielded multiple clusters for all of the AP score groups including the scores of 1 and 5. In some cases members of a given AP score group were divided into as many five or six clusters. For any given cluster, however, only two or three AP groups were included in the cluster. These results were harder to interpret than the seven-cluster solution. Inspection of the significant differences for the clusters did reveal, however, that the clusters labeled as 3s did not differ from one another.

TABLE 37

Biology 1995: Descriptive Statistics of the Biology Outcome Measures

	Mean AP Score Within Cluster	*Frequency*	*Mean*	*Standard Deviation*
BIO 303 Grades	1.0	1	2.00	
	1.8	7	2.86	1.07
	2.6	6	2.67	1.03
	3.1	15	3.00	1.00
	3.8	13	2.77	1.24
	4.4	28	4.00	0.00
	5.0	12	4.00	0.00
Other Biology Hours Taken	1.0	2	4.00	2.83
	1.8	10	4.40	1.90
	2.6	8	2.38	0.52
	3.1	7	3.00	1.73
	3.8	4	2.00	0.00
	4.4	7	2.14	0.38
	5.0	4	2.25	0.50
GPAs in Other Biology Classes	1.0	2	4.00	0.00
	1.8	10	3.17	0.88
	2.6	8	3.00	0.53
	3.1	7	3.86	0.38
	3.8	4	3.00	0.00
	4.4	7	3.43	0.79
	5.0	4	3.75	0.50

TABLE 38

Biology 1996: Descriptive Statistics of the Biology Outcome Measures

	Mean AP Score Within Cluster	*Frequency*	*Mean*	*Standard Deviation*
BIO 303 Grades	1.0	7	2.14	0.69
	1.9	20	2.30	0.92
	2.8	27	2.70	1.30
	2.9	31	3.06	1.00
	4.1	37	3.51	0.80
	4.2	71	3.58	0.84
	5.0	38	4.00	0.00
Other Biology Hours Taken	1.0	19	5.21	2.37
	1.9	25	4.92	2.58
	2.8	24	3.33	2.08
	2.9	21	3.48	2.23
	4.1	18	3.94	3.84
	4.2	26	3.65	2.74
	5.0	21	3.71	2.99
GPAs in Other Biology Classes	1.0	19	2.37	0.81
	1.9	25	2.86	1.01
	2.8	24	2.97	0.99
	2.9	21	2.63	0.94
	4.1	18	3.49	0.78
	4.2	26	3.37	0.69
	5.0	21	3.64	0.55

BIO 303 Grades: 1.0 > 1.0 – 2.9; 4.1, 4.2 > 1.0 – 2.8; 2.9 > 1.9 Other GPAs: 5.0 > 1.0, 1.9, 2.9; 4.1, 4.2 > 1.0, 2.9

TABLE 39

Biology 1997: Descriptive Statistics of the Biology Outcome Measures

	Mean AP Score Within Cluster	*Frequency*	*Mean*	*Standard Deviation*
BIO 303 Grades	1.1	12	1.83	1.11
	1.9	13	3.00	0.71
	2.8	24	2.29	1.49
	3.0	35	2.69	0.87
	4.2(a)	32	3.44	1.13
	4.2(b)	49	3.63	0.78
	5.0	70	4.00	0.00
Other Biology Hours Taken	1.1	31	5.10	2.21
	1.9	29	4.55	2.69
	2.8	16	4.13	3.86
	3.0	36	4.64	3.08
	4.2(a)	23	5.52	4.05
	4.2(b)	24	5.42	3.55
	5.0	32	4.97	3.24
GPAs in Other Biology Classes	1.1	31	2.52	1.31
	1.9	29	2.63	1.14
	2.8	16	3.00	0.89
	3.0	36	2.98	0.88
	4.2(a)	23	3.10	0.93
	4.2(b)	24	3.35	0.94
	5.0	32	3.43	0.94

BIO 303 Grades: 5.0 > 1.1 – 4.2(a); 4.2(a), 4.2(b) > 1.1, 2.8, 3.0; 2.8, 3.0 > 1.1 Other GPAs: 5.0 > 1.1, 1.9; 4.2 > 1.1

TABLE 40

Biology 1998: Descriptive Statistics of the Biology Outcome Measures

	Mean AP Score Within Cluster	*Frequency*	*Mean*	*Standard Deviation*
BIO 303 Grades	1.0	14	2.43	1.16
	1.7	16	3.06	0.77
	2.4	30	2.67	0.84
	3.3	32	2.91	1.25
	3.6	29	2.93	1.19
	4.6	51	3.82	0.71
	5.0	46	4.00	0.00
Other Biology Hours Taken	1.0	23	4.30	1.87
	1.7	29	5.07	2.14
	2.4	33	3.61	1.34
	3.3	34	4.38	3.14
	3.6	21	4.00	2.61
	4.6	20	4.15	2.01
	5.0	21	3.76	2.32
GPAs in Other Biology Classes	1.0	23	2.87	0.93
	1.7	29	2.97	0.96
	2.4	33	2.82	0.88
	3.3	34	3.03	1.03
	3.6	21	2.97	1.15
	4.6	20	3.55	0.78
	5.0	21	3.32	0.64

BIO 303 Grades: 4.6, 5.0 > 1.0 – 3.6

TABLE 41

Biology 1999: Descriptive Statistics of the Biology Outcome Measures

	Mean AP Score Within Cluster	*Frequency*	*Mean*	*Standard Deviation*
BIO 303 Grades	1.1	4	2.00	1.41
	2.1	9	2.44	1.13
	3.2	16	2.25	1.61
	3.6	12	2.50	1.17
	4.4	27	3.70	0.54
	4.5	22	3.64	0.66
	5.0	41	4.00	0.00
Other Biology Hours Taken	1.1	20	3.95	2.09
	2.1	36	4.25	1.36
	3.2	12	3.50	1.45
	3.6	17	3.47	1.70
	4.4	15	3.07	1.16
	4.5	15	3.60	1.40
	5.0	20	4.10	1.89
GPAs in Other Biology Classes	1.1	20	2.67	0.99
	2.1	36	2.72	0.86
	3.2	12	2.59	1.20
	3.6	17	2.69	1.05
	4.4	15	2.64	1.48
	4.5	15	3.02	0.90
	5.0	20	3.69	0.56

BIO 303 Grades: 4.4, 4.5, 5.0 > 1.1 – 3.6 Other GPAs: 5.0 > 1.1 – 3.6

Phase I: Latent Class Analyses

The University of Texas at Austin data sets were used for the latent class analyses. The premise of each of the latent class analyses is that examinees that have earned an AP score of 3 do not reflect a homogeneous group and that this heterogeneity can be determined on the basis of a latent class analysis (LCA). To perform the LCA only examinees that had earned an AP score of 3 were used. In addition, the LCA omitted responses were treated as incorrect and responses only to the multiple-choice questions were used in the analysis. Limitation of the statistical procedure prevented the inclusion of the constructed responses in the analyses.

As can be seen in Table 2, some of the administration dates had a very small number of examinees. This fact coupled with the number of items on the various subject examinations prevented a complete LCA. Specifically, there were insufficient numbers of examinees available to analyze any of the Biology administrations. Similarly, it was not possible to perform a LCA for the 1999 administration in the English Language and Composition Examination or for any of the Calculus AB Examination administrations except for the 1997 test administration. Moreover, for none of the administration dates, irrespective of subject area, was it possible to obtain results for a three-class solution.

To assess the fit of the LC models the Likelihood ratio (G^2), Akaike Information Criterion (AIC) and Bayesian information criterion (BIC) statistics were calculated. G^2, AIC, and BIC are traditional measures for assessing model-data fit in LCA. While all three measures are based on the likelihood function, the AIC allows for a comparison of two models for the same data and has a tendency to favor the model that would exhibit the smallest decrease in likelihood if the model were cross-validated on a new sample. BIC is similar to AIC, but takes into account the sample size. When compared to AIC, BIC tends to select models that are less complex. It should be noted that because of the size of the sample relative to the number of parameters estimated the G^2 statistic may not be very useful. Of the two information-based indices AIC may be preferred because of the relatively small sample sizes involved in this study.

Based on the G^2, AIC, and BIC statistics that are shown in Table 42 for the Calculus administration that could be analyzed, the two-class solution did not appear to provide a meaningfully better fit than the one-class solution. That is, for the Calculus 1997 administration a one-class solution appears to be the preferred solution.

For the four English Language and Composition Examination administrations, the G^2, AIC and BIC statistics showed that the two-class solution provided a better fit than did the one-class solution. It appears that for the English Language and Composition Examination a two-class solution might be considered preferable to a one-class solution. The latent class proportions were 0.223 and 0.777 (1995), 0.132 and 0.868 (1996), 0.633 and 0.367 (1997), and 0.177 and 0.823 (1998). Clearly, for the 1996 and 1998 administrations one of the classes was quite small.

For the four administrations of the English Language and Composition Examination, the interpretation of the two latent classes was difficult given the available variables. For none of the administrations did the two classes differ significantly from one another in terms of either the AP composite score or the AP multiple-choice score.

For the four administrations of the English Language and Composition Examination, independent sample t-tests were conducted to see if the two classes were different in terms of other hours taken in the subject area (HOURS), GPA in other classes taken in the subject area (GPA), average SAT Verbal score (SAT-V), SAT Total score (SAT T), and high school rank (HSR). Table 43 presents the descriptive statistics for each of these outcome measures for each of the four test administration dates.

TABLE 42

Fit Analysis Results

AP Exam	*Test Date*	*Number of Latent Classes*	*G^2*	*AIC*	*BIC*
Calculus	1997	1	6711.97	6793.97	6912.48
		2	6625.30	6789.30	7029.29
English Language and Composition	1995	1	18955.14	19065.14	19266.91
		2	18336.38	18556.38	18963.66
	1996	1	28335.88	28441.88	28660.52
		2	27882.54	28094.54	28536.02
	1997	1	38230.62	38342.62	38582.11
		2	37719.98	37943.98	38427.31
	1998	1	40893.29	41005.29	41250.92
		2	40153.32	40377.32	40873.06

TABLE 43

Descriptive Statistics of the Outcome Measures for the Latent Classes for Each Test Administration Year of the English Language and Composition Examination

Test Date	*Outcome Measure*	*Latent Class*	*N*	*Mean*	*SD*
1995	Hours	1	27	6.39	7.82
		2	60	7.70	7.42
	GPA	1	27	3.67	0.45
		2	60	3.40	0.76
	SAT-V	1	70	640.00	95.95
		2	240	630.17	106.46
	SAT T	1	70	1298.86	106.23
		2	240	1292.63	101.92
	HSR	1	70	89.38	11.11
		2	240	88.51	9.89
1996	Hours	1	12	5.75	5.03
		2	93	7.40	7.98
	GPA	1	12	2.93	1.00
		2	93	3.33	0.83
	SAT-V	1	64	612.03	122.49
		2	431	617.33	97.07
	SAT T	1	64	1283.59	102.98
		2	431	1265.92	107.54
	HSR	1	64	87.03	12.39
		2	431	87.12	11.93
1997	Hours	1	78	5.96	5.79
		2	50	5.10	5.40
	GPA	1	78	3.35	0.85
		2	50	3.37	0.84
	SAT-V	1	365	607.75	115.80
		2	210	586.95	133.97
	SAT T	1	365	1263.67	105.59
		2	210	1251.43	93.30
	HSR	1	365	89.92	9.26
		2	210	85.40	11.46
1998	Hours	1	17	4.06	2.59
		2	67	4.52	3.15
	GPA	1	17	3.35	0.58
		2	67	3.32	0.81
	SAT-V	1	115	599.74	139.70
		2	528	608.31	97.96
	SAT T	1	115	1276.61	96.29
		2	528	1253.83	108.25
	HSR	1	115	87.56	10.63
		2	528	88.93	10.04

None of the dependent variables were found to differ across the latent classes for the 1995 and 1996 test administrations. For the 1997 test administration the two classes differed only in terms of high school rank ($t = 4.871$, $p < .001$). SAT Total score was the only variable to be significantly different across latent classes for the 1999 test administration ($t = 2.084$, $p < .05$).

In summary, the LCAs of the examinees that earned an AP score of 3 seem to indicate that *if* there are two classes they do not differ from one another in a consistent fashion across test administrations. The LCAs were hampered by insufficient sample sizes relative to the number of parameters to be estimated as well as by large test lengths that introduced computational difficulties. Because of this latter factor it was not possible

to perform some of the analyses (e.g., three-class analyses) nor could certain data sets even be analyzed.

Phase II Analyses

Table 44 presents the means and standard deviations of the English outcome measures for each of the four University of Texas at Austin groups used to address research questions 4–6 for four entering classes. The AP-CR group consists of AP students who earned credit by examination for E 306 and then took E 316K. The AP-Class group included AP students who did not earn credit by examination and who took both courses in the sequence. The Non-AP group was matched to the AP-Class group, while the Concurrent group was enrolled in a college-level course equivalent to the E 306 course while they were still in high school. Students who earned credit by examination for both the first and second English courses in the sequence were not included as a comparison group; the group that earned credit by examination for both courses varied in size from 254 in 1999 to 305 in 1997.

ANOVAs were conducted to determine if the four comparison groups differed significantly in the average grades earned in E 316K, the number of other English semester hours taken, and the GPAs earned in the other English classes. For the 1996 entering class, the AP-CR group earned significantly higher average grades than did the Non-AP group ($F = 3.20$, $p < .03$). The AP-CR group also earned significantly higher average grades than did the AP-Class group and the Non-AP group in 1998 ($F = 3.91$, $p < .009$). The only other significant difference that was found in the table was for the average GPAs in other English courses in 1997 ($F = 3.80$, $p < .02$); the AP-CR group had a significantly higher GPA on average than did the students in the concurrent enrollment group. It should be noted that in all cases the AP-CR group had the highest average GPAs in the second course and in subsequent courses in English.

Table 45 presents the same information as Table 44 for the mathematics outcome measures. For all four entering classes, the AP-CR group earned higher grades on average in M 408D than did the students in the AP-Class group and the Non-AP group (1996, $F = 9.73$, $p < .0001$; 1997, $F = 12.21$, $p < .0001$; 1998, $F = 10.22$, $p < .0001$, and 1999, $F = 24.92$, $p < .0001$). Nothing can be said about concurrently enrolled students because there was only one student in this group. In 1997, students in the AP-CR group took on average significantly more hours in mathematics than did students in the Non-AP group ($F = 3.44$, $p < .04$). In 1998 and 1999, the AP-CR group also took on average more hours in mathematics than did the AP-Class group or the Non-AP group (1998, $F = 6.80$, $p < .002$; 1999, $F = 33.30$, $p < .0001$).

The descriptive statistics of the biology outcome measures for the four comparison groups in each of the four entering class are presented in Table 46. The only significant difference in the table was found for the average grades in BIO 303 in 1997 ($F = 3.15$, $p < .03$). The AP-CR group earned a significantly lower average grade in BIO 303 than did the Non-AP group. This finding was not replicated in the other years even though the mean differences were in the same direction.

TABLE 44

Descriptive Statistics for the English Outcome Measures for Four of the University of Texas at Austin Groups

Academic Year	*Group*	*E 316K Grades*			*Other English Hours Taken*			*GPAs in Other English Classes*		
		Mean	*SD*	*Freq.*	*Mean*	*SD*	*Freq.*	*Mean*	*SD*	*Freq.*
1996	AP-CR	3.30	0.86	151	6.94	7.83	48	3.52	0.58	48
	AP-Class	3.01	0.89	81	8.08	9.21	26	3.26	0.64	26
	Non-AP	3.01	0.99	142	5.79	6.80	29	3.26	0.69	29
	Concurrent	3.19	0.92	162	8.47	8.33	45	3.11	1.00	45
1997	AP-CR	3.22	0.89	212	6.50	7.30	48	3.34	0.90	48
	AP-Class	2.96	0.96	82	6.92	6.50	26	3.38	0.66	26
	Non-AP	3.11	0.81	202	5.40	5.08	55	3.26	0.94	55
	Concurrent	3.05	0.94	266	7.65	6.76	40	2.77	1.15	40
1998	AP-CR	3.22	0.93	165	5.50	4.56	48	3.43	0.74	48
	AP-Class	2.86	0.89	83	4.00	1.78	18	3.15	0.57	18
	Non-AP	2.94	0.95	147	5.05	3.47	19	3.02	1.03	19
	Concurrent	3.03	0.91	241	4.77	2.55	39	3.11	0.75	39
1999	AP-CR	3.24	0.81	83	3.25	0.87	12	3.33	0.89	12
	AP-Class	2.89	0.83	28	3.00	0.00	4	3.00	0.00	4
	Non-AP	3.04	0.84	70			0			0
	Concurrent	3.20	0.88	159	3.88	1.41	17	3.24	0.53	17

TABLE 45

Descriptive Statistics for the Mathematics Outcome Measures for Four of the University of Texas at Austin Groups

Academic Year	*Group*	*M 408D Grades*			*Other Math Hours Taken*			*GPAs in Other Math Classes*		
		Mean	*SD*	*Freq.*	*Mean*	*SD*	*Freq.*	*Mean*	*SD*	*Freq.*
1996	AP-CR	2.86	1.18	360	9.13	7.58	267	2.88	1.04	267
	AP-Class	2.53	1.13	126	7.86	5.79	91	2.75	0.91	91
	Non-AP	2.47	1.30	329	9.26	7.31	232	2.82	1.03	232
	Concurrent			0			0			0
1997	AP-CR	2.91	1.14	398	8.42	5.92	307	2.87	1.04	307
	AP-Class	2.61	1.18	176	7.64	5.47	129	2.65	1.06	129
	Non-AP	2.48	1.29	350	7.28	4.69	289	2.70	1.02	289
	Concurrent			0			0			0
1998	AP-CR	2.94	1.14	371	7.37	3.96	262	2.82	1.06	262
	AP-Class	2.47	1.43	166	6.10	2.88	114	2.58	1.14	114
	Non-AP	2.64	1.22	330	6.43	3.40	227	2.75	1.03	227
	Concurrent			0			0			0
1999	AP-CR	3.21	1.08	399	5.63	2.40	273	2.88	1.00	273
	AP-Class	2.48	1.29	223	4.12	1.18	127	2.74	1.08	127
	Non-AP	2.60	1.29	327	4.09	1.11	183	2.69	1.13	183
	Concurrent	0.00		1	4.00		1	3.00		1

TABLE 46

Descriptive Statistics for the Biology Outcome Measures for Four of the University of Texas at Austin Groups

Academic Year	*Group*	*BIO 303 Grades*			*Other Biology Hours Taken*			*GPAs in Other Biology Classes*		
		Mean	*SD*	*Freq.*	*Mean*	*SD*	*Freq.*	*Mean*	*SD*	*Freq.*
1996	AP-CR	2.80	1.13	90	3.04	2.37	48	3.15	0.81	48
	AP-Class	2.92	1.02	38	3.04	2.13	25	3.32	0.70	25
	Non-AP	2.99	1.04	91	3.74	3.38	57	3.33	0.71	57
	Concurrent	2.80	0.84	5	2.25	0.50	4	3.00	0.00	4
1997	AP-CR	2.55	1.21	77	5.40	3.72	43	3.03	0.92	43
	AP-Class	2.68	1.09	65	5.72	3.91	46	2.80	0.95	46
	Non-AP	3.06	1.00	80	5.73	3.37	55	3.12	0.86	55
	Concurrent	2.50	0.71	2	5.00		1	3.00		1
1998	AP-CR	2.84	1.09	74	4.50	2.56	46	2.93	1.05	46
	AP-Class	2.85	1.04	71	5.05	3.05	59	2.81	1.03	59
	Non-AP	2.99	1.12	73	4.64	2.59	50	3.22	0.72	50
	Concurrent	2.50	1.05	6	3.20	1.64	5	3.20	0.84	5
1999	AP-CR	2.64	1.24	53	3.33	1.56	30	2.74	1.09	30
	AP-Class	2.65	1.20	46	3.77	1.70	26	2.95	0.82	26
	Non-AP	2.84	1.16	56	3.42	1.67	24	3.09	0.76	24
	Concurrent	2.00	0.82	4	3.00		1	3.33		1

V. Discussion

Two of the three purposes of the present research were to investigate various procedures to identify low and high AP scores of 3 and to determine if the resulting finer gradations of the AP scale would facilitate universities in course placement of AP students. Several procedures were used to distinguish between low and high scores of 3 for five different administrations of three AP Examinations. The fixed percentage methods were the only procedures that yielded finer gradations of the AP scale that could be considered to be comparable across the different test administrations. The results for these fixed percentage methods, however, showed no significant differences between the various levels of 3s for any of the academic outcome measures for any of the subject areas investigated in this study. In addition, the 3-AP group did not differ significantly from the 2+ or 4-AP groups. While the cluster analyses yielded clusters that were interpretable, the clusters that were labeled 3 did not differ significantly from one another. Similar to the results for the fixed percentage method, the low 3 clusters did not differ from the high 2 clusters or the low 4 clusters. The latent class analysis also failed to find significant differences between the two latent classes of 3s on any of the academic outcome measures. Collectively, the statistical procedures that were investigated yielded similar results and did not produce sufficient evidence to support the use of a scale that distinguished between low and high grades of 3.

The third purpose of the research was to compare the performance of AP students who earn credit by examination with other relevant groups on a number of academic outcome measures for different entering classes of freshman students. The AP group that earned credit by examination was compared to other AP students who took the first course rather than placing out, a matched sample of non-AP students, and students who were concurrently enrolled in the college class while they were still in high school. The results revealed that the AP students who earned credit by examination performed as well if not better on the three academic outcome measures of grades in the sequent course, number of other hours taken in the subject area, and the GPA in the additional courses in the subject area. The students who were concurrently enrolled in college and high school classes at the same time did not differ significantly from any of the other student groups except for their GPA in other English courses. This finding, however, was not replicated in three of the four freshman classes.

VI. Summary of Implications

The primary goals of the present investigation were to assess the validity of grades of 3 on AP Examinations and compare the performance of AP students to other relevant student groups. Phase I of the study investigated three different statistical procedures that could be used to identify "high" and "low" grades of 3 using data from five administrations of three AP Examinations (English Language and Composition, Biology and Calculus AB). The new grade categories were then applied to University of Texas at Austin samples so that the grade groups could be compared in terms of several academic outcome measures. This procedure was replicated for four entering classes of students. The results showed that performance of the low and high 3-grade groups did not differ significantly from one another or the middle 3-grade group. While the results of Phase I demonstrated that finer gradations of grades of 3 could be developed, there was insufficient evidence to warrant changing the current 1–5 scale to a scale that would distinguish between low and high grades of 3.

Phase II of the study compared two AP student groups and two non-AP student groups for each of four entering classes at the University of Texas at Austin. One AP group included AP students who earned credit by examination for the first course and then took the sequent course in the subject area. The other AP group consisted of AP students who did not earn credit by examination and took both courses in the sequence. One of the non-AP groups included students who were matched to the AP student group who took both courses in the sequence. The other non-AP comparison group was comprised of students who were concurrently enrolled in a college class while still in high school. These four groups were compared on three academic outcome measures. In general the results revealed that AP students who earned credit by examination earned equal or higher grades in the sequent course than the other groups. The AP students who earned credit by examination also took as many or more hours in the subject area and had equal or higher grades in additional courses in the subject area than the other groups.

References

Casserly, P. L. (1986). *Advanced placement revisited*. (College Board Report No. 86-6). New York, NY: College Entrance Examination Board.

Dorans, N. J. (1999). *Correspondences between ACT and SAT I scores*. (College Board Report No. 99-1). New York, NY: College Entrance Examination Board.

Koch, W. R., Fitzpatrick, S. J., Triscari, R. S., Mahoney, S. S., and Cope, J. E. (1988). *The Advanced Placement Program: Student attitudes, academic performance, and institutional policies*. Austin: The University of Texas at Austin, Measurement and Evaluation Center.

Morgan, R., & Crone, C. (1993). *Advanced Placement examinees at the University of California: An examination of the freshman year courses and grades of examinees in Biology, Calculus, and Chemistry*. (ETS Statistical Report 93-210). Princeton, NJ: Educational Testing Service.

Morgan, R., & Maneckshana, B. (2000). *AP students in college: An investigation of their course-taking patterns and college majors*. (ETS Statistical Report 2000-09). Princeton, NJ: Educational Testing Service.

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

Appendix: Results of the Eleven-Cluster Solutions

TABLE A.1

English 1995: Eleven National Clusters Against AP Scores

Mean AP Score Within Cluster		*AP Score* 1	2	3	4	5	*Total*
1.0	Number	2,257					2,257
	Column %	42.5%					4.4%
1.5	Number	1,975	2,274				4,249
	Column %	37.2%	12.0%				8.3%
1.7	Number	1,080	2,892				3,972
	Column %	20.3%	15.3%				7.7%
2.0	Number		6,217				6,217
	Column %		32.8%				12.1%
2.3	Number		3,426	1,461			4,887
	Column %		18.1%	9.7%			9.5%
2.4	Number		2,878	1,983	1	1	4,863
	Column %		15.2%	13.2%	0.0%	0.0%	9.5%
2.6	Number	1	864	1,233	14		2,112
	Column %	0.0%	4.6%	8.2%	0.2%		4.1%
3.0	Number		375	6,970	224		7,569
	Column %		2.0%	46.3%	2.6%		14.8%
3.7(a)	Number			2,021	3,702	128	5,851
	Column %			13.4%	43.7%	3.6%	11.4%
3.7(b)	Number			1,393	3,080	89	4,562
	Column %			9.2%	36.4%	2.5%	8.9%
4.7	Number				1,450	3,314	4,764
	Column %				17.1%	93.8%	9.3%
Total	Number	5,313	18,926	15,061	8,471	3,532	51,303

TABLE A.2

English 1996: Eleven National Clusters Against AP Scores

Mean AP Score Within Cluster		*AP Score* 1	2	3	4	5	*Total*
1.0	Number	2,343					2,343
	Column %	56.2%					4.0%
1.7	Number	1,192	2,523				3,715
	Column %	28.6%	13.8%				6.3%
1.9	Number	629	3,942				4,571
	Column %	15.1%	21.5%				7.7%
2.1	Number		7,126	526			7,652
	Column %		38.9%	2.6%			13.0%
2.3	Number	3	2,602	1,166			3,771
	Column %	0.1%	14.2%	5.8%			6.4%
2.6	Number		2,115	3,628			5,743
	Column %		11.6%	18.1%			9.7%
3.0	Number			9,057	294		9,351
	Column %			45.1%	2.4%		15.9%
3.3	Number			3,649	1,645		5,294
	Column %			18.2%	13.4%		9.0%
3.7	Number			2,053	4,272	15	6,340
	Column %			10.2%	34.9%	0.4%	10.7%
4.2	Number				5,640	1,154	6,794
	Column %				46.1%	27.6%	11.5%
4.9	Number				393	3,014	3,407
	Column %				3.2%	72.1%	5.8%
Total	Number	4,167	18,308	20,079	12,244	4,183	58,981

TABLE A.3

English 1997: Eleven National Clusters Against AP Scores

Mean AP Score Within Cluster		*AP Score* 1	2	3	4	5	*Total*
1.1	Number	3,147	184				3,331
	Column %	79.2%	0.9%				5.0%
1.9(a)	Number	189	3,498				3,687
	Column %	4.8%	18.1%				5.5%
1.9(b)	Number	635	5,966				6,601
	Column %	16.0%	30.8%				9.9%
2.3	Number		6,174	2,375			8,549
	Column %		31.9%	9.9%			12.8%
2.6	Number		2,522	3,898			6,420
	Column %		13.0%	16.2%			9.6%
2.8	Number		1,028	3,801			4,829
	Column %		5.3%	15.8%			7.2%
3.0	Number			9,138	364		9,502
	Column %			38.0%	2.9%		14.2%
3.6(a)	Number			2,108	3,073	75	5,256
	Column %			8.8%	24.5%	1.1%	7.8%
3.6(b)	Number			2,738	3,981	83	6,802
	Column %			11.4%	31.8%	1.2%	10.2%
4.3	Number				5,043	2,547	7,590
	Column %				40.2%	36.3%	11.3%
5.0	Number				75	4,319	4,394
	Column %				0.6%	61.5%	6.6%
Total	Number	3,971	19,372	24,058	12,536	7,024	66,961

TABLE A.4

English 1998: Eleven National Clusters Against AP Scores

Mean AP Score Within Cluster		*AP Score* 1	2	3	4	5	*Total*
1.0	Number	2,268					2,268
	Column %	45.7%					2.9%
1.5	Number	2,368	2,738				5,106
	Column %	47.7%	12.0%				6.4%
1.9	Number	301	3,583				3,884
	Column %	6.1%	15.7%				4.9%
2.0	Number	25	5,882				5,907
	Column %	0.5%	25.7%				7.4%
2.2	Number		7,991	1,819			9,810
	Column %		34.9%	6.8%			12.3%
2.7	Number		2,285	5,429			7,714
	Column %		10.0%	20.2%			9.7%
3.1(a)	Number			10,877	630		11,507
	Column %			40.4%	3.7%		14.5%
3.1(b)	Number		410	6,519	1,066		7,995
	Column %		1.8%	24.2%	6.2%		10.1%
3.7	Number			2,291	5,702	37	8,030
	Column %			8.5%	33.2%	0.5%	10.1%
4.2	Number				9,186	1,951	11,137
	Column %				53.6%	26.0%	14.0%
4.9	Number				568	5520	6,088
	Column %				3.3%	73.5%	7.7%
Total	Number	4,962	22,889	26,935	17,152	7,508	79,446

TABLE A.5

English 1999: Eleven National Clusters Against AP Scores

Mean AP Score Within Cluster		*AP Score* 1	2	3	4	5	*Total*
1.2	Number	3,736	705				4,441
	Column %	82.6%	2.2%				4.6%
1.9	Number	739	7,693				8,432
	Column %	16.3%	24.1%				8.7%
2.0	Number	46	5,325				5,371
	Column %	1.0%	16.7%				5.6%
2.1	Number		8,131	642			8,773
	Column %		25.4%	1.9%			9.1%
2.3	Number		9,312	3,986			13,298
	Column %		29.1%	11.8%			13.8%
2.9	Number		760	6,646	208		7,614
	Column %		2.4%	19.6%	1.2%		7.9%
3.0	Number		45	8,413	116	1	8,575
	Column %		0.1%	24.8%	0.7%	0.0%	8.9%
3.1	Number			12,789	1,045		13,834
	Column %			37.8%	5.9%		14.3%
4.0	Number			1,032	7,956	798	9,786
	Column %			3.0%	45.1%	9.2%	10.1%
4.1	Number			353	8,078	1,515	9,946
	Column %			1.0%	45.8%	17.4%	10.3%
5.0	Number				251	6,382	6,633
	Column %				1.4%	73.4	6.9%
Total	Number	4,521	31,971	33,861	17,654	8,696	96,703

TABLE A.6

English 1995: Descriptive Statistics of the English Outcome Measures

	Mean AP Score Within Cluster	*Frequency*	*Mean*	*Standard Deviation*
E 316K Grades	1.0	21	3.52	0.60
	1.5	56	3.29	0.71
	1.7	35	3.06	0.91
	2.0	51	3.29	0.78
	2.3	42	3.33	0.69
	2.4	37	3.08	0.80
	2.6	18	3.61	0.50
	3.0	41	3.54	0.55
	3.7(a)	48	3.79	0.41
	3.7(b)	36	3.75	0.44
	4.7	59	3.80	0.48
Other English Hours Taken	1.0	3	7.00	6.93
	1.5	18	4.50	5.66
	1.7	22	7.36	7.21
	2.0	35	4.97	4.88
	2.3	21	10.19	10.34
	2.4	19	6.79	6.31
	2.6	13	13.62	10.44
	3.7(a)	38	7.34	7.24
	3.7(b)	43	10.81	11.32
	3.7	32	10.50	12.61
	4.7	42	16.71	12.24
GPAs in Other English Classes	1.0	3	3.20	1.06
	1.5	18	3.57	0.50
	1.7	22	3.06	0.99
	2.0	35	3.30	0.89
	2.3	21	3.19	0.61
	2.4	19	3.63	0.57
	2.6	13	3.82	0.24
	3.0	38	3.48	0.84
	3.7(a)	43	3.58	0.70
	3.7(b)	32	3.53	1.00
	4.7	42	3.69	0.45

E 316K Grades ($F = 7.35$, $p < .001$): 3.7(b), 4.7 > 1.5 – 2.4
3.7(a) > 1.5 – 2.0, 2.4
3.0 > 1.7

Other Hours ($F = 4.52$, $p < .001$): 4.7 > 1.5 – 2.0, 2.4, 3.0

TABLE A.7

English 1996: Descriptive Statistics of the English Outcome Measures

	Mean AP Score Within Cluster	*Frequency*	*Mean*	*Standard Deviation*
E 316K Grades	1.0	19	3.05	0.62
	1.7	39	3.13	0.98
	1.9	57	3.25	0.76
	2.1	95	3.21	0.86
	2.3	57	3.28	0.73
	2.6	82	3.24	0.75
	3.0	104	3.36	0.87
	3.3	51	3.57	0.64
	3.7	59	3.56	0.57
	4.2	63	3.70	0.66
	4.9	52	3.85	0.41
Other English Hours Taken	1.0	7	5.57	6.80
	1.7	6	7.50	9.63
	1.9	21	8.14	7.95
	2.1	38	6.16	6.65
	2.3	26	3.81	8.03
	2.3	26	3.81	8.03
	3.0	63	8.00	7.94
	3.3	45	10.07	8.51
	3.7	46	9.13	7.92
	4.2	52	11.31	8.94
	4.9	52	16.44	11.19
GPAs in Other English Classes	1.0	7	3.55	0.56
	1.7	6	3.47	0.45
	1.9	21	3.31	0.95
	2.1	38	3.38	0.68
	2.3	26	3.01	1.16
	2.6	30	3.54	0.49
	3.0	63	3.39	0.69
	3.3	45	3.43	1.10
	3.7	46	3.40	0.84
	4.2	52	3.41	0.92
	4.9	52	3.77	0.40

E 316K Grades ($F = 5.68$, $p < .001$): 4.9 > 1.0 – 3.0
4.2 > 1.0 – 2.1, 2.6

Other Hours ($F = 5.66$, $p < .001$): 4.9 > 1.9 – 3.7

TABLE A.8

English 1997: Descriptive Statistics of the English Outcome Measures

	Mean AP Score Within Cluster	*Frequency*	*Mean*	*Standard Deviation*
E 316K Grades	1.1	32	3.06	0.84
	1.9(a)	36	2.81	1.14
	1.9(b)	50	3.00	0.83
	2.3	92	3.17	0.66
	2.6	40	3.35	0.62
	2.8	41	3.07	1.10
	3.0	79	3.52	0.57
	3.6(a)	46	3.41	0.93
	3.6(b)	52	3.29	0.98
	4.3	70	3.69	0.55
	5.0	46	3.89	0.31
Other English Hours Taken	1.1	4	4.50	3.00
	1.9(a)	15	4.20	1.90
	1.9(b)	18	4.33	2.57
	2.3	38	4.82	3.98
	2.6	26	5.54	4.62
	2.8	36	6.58	6.37
	3.0	56	6.11	5.75
	3.6(a)	24	8.63	8.07
	3.6(b)	57	5.95	5.05
	4.3	62	9.73	7.01
	5.0	34	11.47	7.29
GPAs in Other English Classes	1.1	4	3.33	0.94
	1.9(a)	15	3.27	0.68
	1.9(b)	18	3.23	0.55
	2.3	38	3.16	0.97
	2.6	26	3.63	0.46
	2.8	36	3.09	1.01
	3.0	56	3.48	0.76
	3.6(a)	24	3.67	0.57
	3.6(b)	57	3.63	0.56
	4.3	62	3.55	0.78
	5.0	34	3.75	0.46

E 316K Grades ($F = 8.20$, $p < .001$): 5.0 > 1.1 – 2.8
4.3 > 1.1 – 2.3, 2.8
3.0 > 1.9(a), 1.9(b)
3.6 > 1.9(a)

Other Hours ($F = 5.36$, $p < .001$): 5.0 > 1.1 – 2.8, 3.6
4.3 > 1.9(b), 2.3, 3.0, 3.6

Other GPAs ($F = 3.12$, $p < .001$): 5.0, 3.6 > 2.6
5.0 > 2.3

TABLE A.9

English 1998: Descriptive Statistics of the English Outcome Measures

	Mean AP Score Within Cluster	*Frequency*	*Mean*	*Standard Deviation*
E 316K Grades	1.0	13	3.23	0.83
	1.5	45	3.09	0.70
	1.9	19	3.00	0.58
	2.0	31	3.06	0.96
	2.2	69	3.35	0.76
	2.7	37	3.35	0.68
	3.1(a)	77	3.35	0.85
	3.1(b)	51	3.37	0.72
	3.7	45	3.62	0.53
	4.2	86	3.78	0.47
	4.9	81	3.81	0.50
Other English Hours Taken	1.0	2	3.00	0.00
	1.5	8	3.00	0.00
	1.9	8	3.38	1.06
	2.0	10	3.30	0.95
	2.2	20	3.00	0.00
	2.7	24	4.25	3.05
	3.1(a)	46	4.89	3.54
	3.1(b)	27	5.44	3.23
	3.7	36	5.75	4.08
	4.2	57	7.47	4.64
	4.9	51	8.35	4.26
GPAs in Other English Classes	1.0	2	2.50	0.71
	1.5	8	3.25	0.71
	1.9	8	3.00	0.93
	2.0	10	3.15	0.58
	2.2	20	3.30	0.80
	2.7	24	3.20	1.01
	3.1(a)	46	3.48	0.65
	3.1(b)	27	3.60	0.54
	3.7	36	3.75	0.39
	4.2	57	3.75	0.50
	4.9	51	3.73	0.46

E 316K Grades (F= 8.26 p < .001): 4.9 > 1.9 – 3.1
4.2 > 1.5 – 2.2, 3.1(a), 3.1(b)
3.7 > 1.5, 1.9, 2.0
3.6 > 1.9(a)

Other Hours (F = 6.92, p < .001): 4.9 > 1.5 – 3.7
4.2 > 1.5, 2.0, 2.7, 3.1(a)

Other GPAs (F = 3.12, p < .001): 3.7 – 4.9 > 2.7

TABLE A.10

English 1999: Descriptive Statistics of the English Outcome Measures

	Mean AP Score Within Cluster	*Frequency*	*Mean*	*Standard Deviation*
E 316K Grades	1.2	1	4.00	.
	1.9	7	2.86	0.90
	2.0	4	2.25	1.50
	2.1	6	3.00	0.63
	2.3	8	3.25	0.89
	2.9	8	3.50	0.76
	3.0	6	3.33	0.52
	3.1	8	3.63	0.52
	4.0	9	3.89	0.33
	4.1	12	3.33	0.49
	5.0	7	4.00	0.00
Other English Hours Taken	1.2	1	3.00	
	1.9	0		
	2.0	1	3.00	
	2.1	0		
	2.3	6	3.00	0.00
	2.9	3	5.00	3.46
	3.0	8	3.00	0.00
	3.1	15	4.20	3.17
	4.0	8	6.38	4.66
	4.1	6	5.50	3.99
	5.0	5	7.80	5.45
GPAs in Other English Classes	1.2	1	4.00	
	1.9	0		
	2.0	1	3.00	
	2.1	0		
	2.3	6	3.67	0.52
	2.9	3	2.89	1.64
	3.0	8	3.12	0.83
	3.1	15	3.43	1.05
	4.0	8	3.75	0.38
	4.1	6	3.19	0.70
	5.0	5	3.78	0.30

TABLE A.11

Calculus 1995: Eleven National Clusters Against AP Scores

Mean AP Score Within Cluster		*AP Score* 1	2	3	4	5	*Total*
1.0	Number	10,907					10,907
	Column %	46.9%					10.6%
1.1	Number	8,548	865				9,413
	Column %	36.7%	4.6%				9.2%
1.6	Number	3,821	5,591	1			9,413
	Column %	16.4%	29.5%	0.0%			9.2%
2.2	Number		9,228	1,934			11,162
	Column %		48.7%	6.7%			10.9%
2.7	Number		3,266	7,624			10,890
	Column %		17.2%	26.4%			10.6%
3.0	Number			11,442	1		11,443
	Column %			39.2%	0.0%		11.1%
3.4	Number			6,604	3,825		10,429
	Column %			22.9%	19.5%		10.1%
3.9	Number			1,271	8,107		9,378
	Column %			4.4%	41.4%		9.1%
4.3	Number				5,528	2,309	7,837
	Column %				28.2%	19.0%	7.6%
4.7	Number				2,129	4,399	6,528
	Column %				10.9%	36.3%	6.3%
5.0	Number					5,420	5,420
	Column %					44.7%	5.3%
Total	Number	23,276	18,950	28,876	19,590	12,128	102,820

TABLE A.12

Calculus 1996: Eleven National Clusters Against AP Scores

Mean AP Score Within Cluster		*AP Score* 1	2	3	4	5	*Total*
1.0	Number	13,083	1				13,084
	Column %	64.5%	0.0%				12.5%
1.6	Number	4,452	5,721	1			10,174
	Column %	21.9%	27.1%	0.0%			9.7%
1.8	Number	2,759	9,213				11,972
	Column %	13.6%	43.7%				11.4%
2.6	Number		5,327	7,607			12,934
	Column %		25.3%	26.8%			12.3%
2.9	Number		684	10,230	22		10,936
	Column %		3.2%	36.1%	0.1%		10.4%
3.1	Number		136	4,916	788		5,840
	Column %		0.6%	17.3%	3.6%		5.6%
3.5	Number			5,538	5,695		11,233
	Column %			19.5%	26.0%		10.7%
4.0	Number			71	8,384	93	8,548
	Column %			0.3%	38.3%	0.7%	8.1%
4.2	Number				6,347	2,092	8,439
	Column %				29.0%	15.6%	8.0%
4.9	Number				635	6,537	7,172
	Column %				2.9%	48.7%	6.8%
5.0	Number					4,708	4,708
	Column %					35.1%	4.5%
Total	Number	20,294	21,082	28,363	21,871	13,430	105,040

TABLE A.13

Calculus 1997: Eleven National Clusters Against AP Scores

Mean AP Score Within Cluster		*AP Score* 1	2	3	4	5	*Total*
1.0	Number	10,235					10,235
	Column %	44.9%					9.2%
1.1	Number	8,880	1,147				10,027
	Column %	39.0%	5.2%				9.0%
1.6	Number	3,680	6,393				10,073
	Column %	16.1%	28.9%				9.1%
2.1	Number		10,779	1,454			12,233
	Column %		48.7%	4.8%			11.0%
2.7	Number		3,803	7,854			11,657
	Column %		17.2%	25.9%			10.5%
3.0	Number			12,565	1		12,566
	Column %			41.4%	0.0%		11.3%
3.4	Number			7,424	4,099		11,523
	Column %			24.5%	18.1%		10.4%
3.9	Number			1,051	9,994	1	11,046
	Column %			3.5%	44.1%	0.0%	9.9%
4.3	Number				6,218	2,568	8,786
	Column %				27.4%	19.4%	7.9%
4.7	Number				2,374	4,903	7,277
	Column %				10.5%	37.0%	6.5%
5.0	Number					5,779	5,779
	Column %					43.6%	5.2%
Total	Number	22,795	22,122	30,348	22,686	13,251	111,202

TABLE A.14

Calculus 1998: Eleven National Clusters Against AP Scores

Mean AP Score Within Cluster		*AP Score* 1	2	3	4	5	*Total*
1.0	Number	9,814					9,814
	Column %	52.0%					8.4%
1.3	Number	6,922	2,601				9,523
	Column %	36.7%	12.5%				8.2%
1.8	Number	2,140	8,163				10,303
	Column %	11.3%	39.4%				8.8%
2.3	Number		6,620	3,335			9,955
	Column %		31.9%	10.7%			8.5%
2.7	Number		3,348	9,259			12,607
	Column %		16.1%	29.6%			10.8%
3.1	Number			8,965	1,528		10,493
	Column %			28.7%	5.6%		9.0%
3.2	Number			9,728	2,855		12,583
	Column %			31.1%	10.5%		10.8%
4.0	Number				12,870		12,870
	Column %				47.5%		11.0%
4.2	Number				7,956	1,928	9,884
	Column %				29.4%	10.4%	8.5%
4.8	Number				1,894	7,846	9,740
	Column %				7.0%	42.4%	8.4%
5.0	Number					8,748	8,748
	Column %					47.2%	7.5%
Total	Number	18,876	20,732	31,287	27,103	18,522	116,520

TABLE A.15

Calculus 1999: Eleven National Clusters Against AP Scores

Mean AP Score Within Cluster		*AP Score*					
		1	*2*	*3*	*4*	*5*	*Total*
1.0	Number	16,223					16,223
	Column %	75.0%					12.9%
1.7	Number	3,873	7,924				11,797
	Column %	17.9%	33.0%				9.4%
1.9	Number	1,538	11,583	66			13,187
	Column %	7.1%	48.3%	0.2%			10.5%
2.7	Number		4,443	8,703			13,146
	Column %		18.5%	27.8%			10.5%
3.0	Number		50	12,515			12,565
	Column %		0.2%	39.9%			10.0%
3.3	Number			8,858	4,211		13,069
	Column %			28.3%	14.8%		10.4%
3.9	Number			1,197	10,364		11,561
	Column %			3.8%	36.3%		9.2%
4.0	Number				10,647	106	10,753
	Column %				37.3%	0.5%	8.5%
4.6	Number				3,312	5,713	9,025
	Column %				11.6%	28.2%	7.2%
5.0	Number					6,819	6,819
	Column %					33.6%	5.4%
5.0	Number				1	7,633	7,634
	Column %				0.0%	37.7%	6.1%
Total	Number	21,634	24,000	31,339	28,535	20,271	125,779

TABLE A.16

Calculus 1995: Descriptive Statistics of the Mathematics Outcome Measures

	Mean AP Score Within Cluster	*Frequency*	*Mean*	*Standard Deviation*
M 408D Grades	1.0	1	2.00	
	1.1	1	3.00	
	1.6	3	1.67	1.53
	2.2	7	2.57	1.27
	2.7	4	2.50	1.29
	3.0	5	3.60	0.55
	3.4	4	3.25	0.96
	3.9	5	3.00	1.22
	4.3	3	2.33	0.58
	4.7	4	4.00	0.00
	5.0	3	4.00	0.00
Other Math Hours Taken	1.0	6	5.00	2.37
	1.1	2	3.50	0.71
	1.6	5	6.40	2.88
	2.2	8	4.38	1.69
	2.7	10	11.70	12.14
	3.0	5	4.60	1.34
	3.4	4	7.25	2.87
	3.9	6	6.67	2.88
	4.3	4	5.50	2.38
	4.7	5	22.60	18.50
	5.0	2	7.00	0.00
GPAs in Other Math Classes	1.0	6	2.00	1.90
	1.1	2	3.00	0.00
	1.6	5	2.50	1.92
	2.2	8	2.89	1.37
	2.7	10	3.03	1.14
	3.0	5	3.60	0.89
	3.4	4	2.78	1.02
	3.9	6	3.11	0.93
	4.3	4	3.48	0.71
	4.7	5	3.71	0.29
	5.0	2	4.00	0.00

Other Hours ($F = 2.43$, $p < .02$): 4.7 > 1.0, 2.2, 3.0

TABLE A.17

Calculus 1996: Descriptive Statistics of the Mathematics Outcome Measures

	Mean AP Score Within Cluster	*Frequency*	*Mean*	*Standard Deviation*
Math 408D Grades	1.0	19	2.00	0.94
	1.6	24	2.46	1.10
	1.8	40	2.53	1.18
	2.6	40	2.42	1.22
	2.9	58	2.34	1.12
	3.1	38	2.13	1.32
	3.5	70	2.79	1.26
	4.0	38	3.11	0.95
	4.2	43	2.91	1.34
	4.9	51	3.16	1.10
	5.0	45	3.44	1.01
Other Math Hours Taken	1.0	59	6.64	4.04
	1.6	61	8.39	5.32
	1.8	57	6.93	5.90
	2.6	61	6.44	4.75
	2.9	74	6.22	3.47
	3.1	46	7.17	5.29
	3.5	79	7.73	6.82
	4.2	43	8.05	6.48
	4.2	43	8.05	6.48
	4.9	59	8.00	7.05
	5.0	45	10.40	9.45
GPAs in Other Math Classes	1.0	59	2.56	0.99
	1.6	61	3.01	0.82
	1.8	57	2.95	0.89
	2.6	61	2.95	1.02
	2.9	74	3.10	0.92
	3.1	46	2.93	1.27
	3.5	79	3.05	1.07
	4.0	38	3.43	0.64
	4.2	43	2.96	1.18
	4.9	59	3.27	0.89
	5.0	45	3.36	0.91

M 408D Grades ($F = 5.81$, $p < .001$): 5.0 > 1.0 – 3.1
4.9 > 1.0, 2.9, 3.1
4.0 > 1.0, 3.0

Other Hours ($F = 2.17$, $p < .02$): 5.0 > 1.0, 2.6, 2.9

Other GPAs ($F = 3.04$, $p < .001$): 5.0, 4.9, 4.0 > 1.0

TABLE A.18

Calculus 1997: Descriptive Statistics of the Mathematics Outcome Measures

	Mean AP Score Within Cluster	*Frequency*	*Mean*	*Standard Deviation*
M 408D Grades	1.0	19	1.68	1.49
	1.1	49	2.16	1.39
	1.6	43	2.72	0.96
	2.1	54	2.70	1.02
	2.7	60	2.83	0.98
	3.0	54	2.61	1.19
	3.4	75	2.80	1.10
	3.9	70	2.80	1.12
	4.3	65	3.14	1.21
	4.7	51	3.53	0.76
	5.0	37	3.38	1.21
Other Math Hours Taken	1.0	59	5.80	2.70
	1.1	75	6.80	4.03
	1.6	57	7.32	5.45
	2.1	67	6.46	3.93
	2.7	67	7.57	6.26
	3.0	71	6.62	3.23
	3.4	90	5.82	2.91
	3.9	72	6.96	4.89
	4.3	75	7.59	5.93
	4.7	63	7.24	4.80
	5.0	36	7.97	5.77
GPAs in Other Math Classes	1.0	59	2.27	1.16
	1.1	75	2.73	0.95
	1.6	57	2.53	1.15
	2.1	67	2.87	1.10
	2.7	67	2.87	1.01
	3.0	71	2.97	0.96
	3.4	90	3.00	1.00
	3.9	72	2.92	1.09
	4.3	75	3.16	0.95
	4.7	63	3.17	1.00
	5.0	36	3.48	0.87

M 408D Grades ($F = 7.41$, $p < .001$): 4.7 > 1.0 – 3.9
4.3 > 1.0, 1.1
1.6 – 2.7, 3.4, 3.9, 5.0 > 1.0

Other GPAs ($F = 5.39$, $p < .001$): 2.1 – 5.0 > 1.0
4.3, 4.7 > 1.6
5.0 > 1.1, 1.6

TABLE A.19

Calculus 1998: Descriptive Statistics of the Mathematics Outcome Measures

	Mean AP Score Within Cluster	*Frequency*	*Mean*	*Standard Deviation*
M 408D Grades	1.0	12	2.58	1.51
	1.3	24	2.50	1.38
	1.8	39	2.05	1.34
	2.3	40	2.17	1.32
	2.7	55	2.35	1.24
	3.1	62	2.98	1.17
	3.2	56	2.82	1.18
	4.0	87	2.89	1.20
	4.2	52	2.98	1.11
	4.8	71	3.11	1.04
	5.0	51	3.45	0.73
Other Math Hours Taken	1.0	57	6.02	2.92
	1.3	45	6.44	2.67
	1.8	54	5.65	2.40
	2.3	48	6.21	2.77
	2.7	64	6.16	3.56
	3.1	72	5.56	2.32
	3.2	68	6.25	4.08
	4.0	85	6.32	3.31
	4.2	51	6.41	3.06
	4.8	60	6.67	3.03
	5.0	47	7.49	5.45
GPA in Other Math Classes	1.0	57	2.74	1.06
	1.3	45	2.64	1.03
	1.8	54	2.93	0.95
	2.3	48	2.74	1.01
	2.7	64	2.74	1.02
	3.1	72	3.08	1.08
	3.2	68	3.00	0.97
	4.0	85	2.96	1.08
	4.2	51	3.01	0.98
	4.8	60	2.89	1.15
	5.0	47	3.58	0.56

M 408D Grades ($F = 6.20$, $p < .001$): 5.0 > 1.3 – 2.3
4.8 > 1.8 – 2.7
3.1, 4.2 > 1.8, 2.3
4.0 > 1.8

Other GPAS ($F = 3.13$, $p < .001$): 5.0 > 1.0, 1.3, 2.3, 2.7, 4.0, 4.8

TABLE A.20

Calculus 1999: Descriptive Statistics of the Mathematics Outcome Measures

	Mean AP Score Within Cluster	*Frequency*	*Mean*	*Standard Deviation*
M 408D Grades	1.0	38	1.61	1.31
	1.7	35	2.23	1.21
	1.9	38	2.84	1.13
	2.7	61	2.56	1.43
	3.0	61	2.75	1.19
	3.3	52	3.15	1.06
	3.9	56	3.07	1.23
	4.0	62	3.19	1.04
	4.6	46	3.00	1.14
	5.0(a)	37	3.54	0.84
	5.0(b)	60	3.70	0.77
Other Math Hours Taken	1.0	86	4.65	2.29
	1.7	60	5.08	2.16
	1.9	51	5.47	2.18
	2.7	64	4.98	1.96
	3.0	64	5.08	2.77
	3.3	56	4.89	1.87
	3.9	55	4.98	1.85
	4.0	62	5.08	2.23
	4.6	58	4.84	1.93
	5.0(a)	32	4.75	1.98
	5.0(b)	49	5.47	1.75
GPAs in Other Math Classes	1.0	86	2.75	1.14
	1.7	60	2.96	0.87
	1.9	51	3.05	0.93
	2.7	64	3.08	0.99
	3.0	64	2.69	1.09
	3.3	56	3.05	0.95
	3.9	55	3.00	1.10
	4.0	62	3.14	0.94
	4.6	58	3.10	0.99
	5.0(a)	32	3.51	0.68
	5.0(b)	49	3.45	0.93

M 408D Grades ($F = 11.75$, $p < .001$): 1.9 – 5.0(b) > 1.0
5.0(a) > 1.7, 2.7, 3.0
5.0(b) > 1.7 – 3.0
4.0 > 1.7

Other GPAs ($F = 3.23$, $p < .001$): 5.0(a), 5.0(b) > 1.0, 3.0

TABLE A.21

Biology 1995: Eleven National Clusters Against AP Scores

Mean AP Score Within Cluster		*AP Score* 1	2	3	4	5	*Total*
1.0	Number	3,758		1			3,759
	Column %	45.7%		0.0%			6.2%
1.4	Number	4,431	2,435				6,866
	Column %	53.9%	17.5%				11.2%
2.0	Number	32	5,538				5,570
	Column %	0.4%	39.7%				9.1%
2.1	Number	1	5,289	862			6,152
	Column %	0.0%	38.0%	5.7%			10.1%
2.9	Number		671	5,847			6,518
	Column %		4.8%	38.6%			10.7%
3.0	Number			6,424	69		6,493
	Column %			42.4%	0.5%		10.6%
3.7	Number			1,817	4,181		5,998
	Column %			12.0%	31.5%		9.8%
4.0	Number			187	5,798		5,985
	Column %			1.2%	43.7%		9.8%
4.5	Number				2,435	2,752	5,187
	Column %				18.4%	26.1%	8.5%
4.8	Number				786	3,678	4,464
	Column %				5.9%	34.9%	7.3%
5.0	Number					4,120	4,120
	Column %					39.1%	6.7%
Total	Number	8,222	13,933	15,138	13,269	10,550	61,112

TABLE A.22

Biology 1996: Eleven National Clusters Against AP Scores

Mean AP Score Within Cluster		*AP Score* 1	2	3	4	5	*Total*
1.0	Number	4,988					4,988
	Column %	53.4%					7.6%
1.4	Number	3,955	3,193				7,148
	Column %	42.3%	22.3%				10.9%
1.9	Number	396	5,571				5,967
	Column %	4.2%	38.8%				9.1%
2.2	Number		4,235	1,361			5,596
	Column %		29.5%	8.9%			8.5%
2.8	Number		1,262	6,052			7,314
	Column %		8.8%	39.4%			11.1%
3.2	Number		79	4,851	1,015		5,945
	Column %		0.6%	31.6%	7.1%		9.1%
3.5	Number			2,991	3,175	1	6,167
	Column %			19.5%	22.2%	0.0%	9.4%
4.0	Number			93	6,628		6,721
	Column %			0.6%	46.3%		10.2%
4.5	Number				2,297	2,629	4,926
	Column %				16.1%	21.4%	7.5%
4.8	Number				1,186	4,673	5,859
	Column %				8.3%	38.0%	8.9%
5.0	Number					4,995	4,995
	Column %					40.6%	7.6%
Total	Number	9,339	14,340	15,348	14,301	12,298	65,626

TABLE A.23

Biology 1997: Eleven National Clusters Against AP Scores

Mean AP Score Within Cluster		*AP Score*					
		1	*2*	*3*	*4*	*5*	*Total*
1.0	Number	4,041					4,041
	Column %	50.0%					5.7%
1.4	Number	3,785	2,450				6,235
	Column %	46.9%	16.5%				8.8%
2.0	Number	249	4,879	251			5,379
	Column %	3.1%	32.9%	1.4%			7.6%
2.2	Number	1	6,967	1,239			8,207
	Column %	0.0%	47.0%	7.0%			11.6%
2.9	Number		498	7,708			8,206
	Column %		3.4%	43.2%			11.6%
3.1	Number		36	6,518	718		7,272
	Column %		0.2%	36.6%	4.6%		10.3%
3.8	Number			992	2,694	89	3,775
	Column %			5.6%	17.1%	0.6%	5.3%
3.9	Number			1,119	6,747		7,866
	Column %			6.3%	42.8%		11.1%
4.3	Number				4,357	2,220	6,577
	Column %				27.7%	15.8%	9.3%
4.8	Number				1,236	5,749	6,985
	Column %				7.8%	40.8%	9.9%
5.0	Number					6,028	6,028
	Column %					42.8%	8.5%
Total	Number	8,076	14,830	17,827	15,752	14,086	70,571

TABLE A.24

Biology 1998: Eleven National Clusters Against AP Scores

Mean AP Score Within Cluster		*AP Score*					
		1	*2*	*3*	*4*	*5*	*Total*
1.0	Number	5,131					5,131
	Column %	42.1%					6.8%
1.3	Number	6,068	2,037				8,105
	Column %	49.8%	11.8%				10.7%
1.8	Number	991	5,367				6,358
	Column %	8.1%	31.1%				8.4%
2.1	Number		6,563	991			7,554
	Column %		38.0%	5.5%			10.0%
2.6	Number		3,304	5,781			9,085
	Column %		19.1%	32.3%			12.0%
3.1	Number			7,142	1,122		8,264
	Column %			39.9%	8.1%		11.0%
3.5	Number			4,000	3,832		7,832
	Column %			22.3%	27.6%		10.4%
4.2	Number				6,606	1,212	7,818
	Column %				47.5%	8.5%	10.4%
4.6	Number				2,333	3,238	5,571
	Column %				16.8%	22.8%	7.4%
5.0(a)	Number					5,594	5,594
	Column %					39.4%	7.4%
5.0(b)	Number					4,153	4,153
	Column %					29.3%	5.5%
Total	Number	12,190	17,271	17,914	13,893	14,197	75,465

TABLE A.25

Biology 1999: Eleven National Clusters Against AP Scores

Mean AP Score Within Cluster		*AP Score*					*Total*
		1	*2*	*3*	*4*	*5*	
1.0	Number	6,968					6,968
	Column %	65.5%					8.6%
1.7	Number	3,642	7,432				11,074
	Column %	34.2%	41.8%				13.6%
2.3	Number	28	5,463	2,416			7,907
	Column %	0.3%	30.7%	12.6%			9.7%
2.5	Number		4,871	4,475			9,346
	Column %		27.4%	23.4%			11.5%
3.2	Number			8,077	1,635		9,712
	Column %			42.2%	9.1%		11.9%
3.6	Number			2,269	3,421	16	5,706
	Column %			11.9%	19.0%	0.1%	7.0%
3.7	Number			1,910	4,384		6,294
	Column %			10.0%	24.4%		7.7%
4.2	Number				6,880	2,194	9,074
	Column %				38.3%	13.7%	11.1%
4.7	Number				1,658	4,082	5,740
	Column %				9.2%	25.6%	7.0%
5.0(a)	Number					5,463	5,463
	Column %					34.2%	6.7%
5.0(b)	Number					4,208	4,208
	Column %					26.4%	5.2%
Total	Number	10,638	17,766	19,147	17,978	15,963	81,492

TABLE A.26

Biology 1995: Descriptive Statistics of the Biology Outcome Measures

	Mean AP Score Within Cluster	*Frequency*	*Mean*	*Standard Deviation*
BIO 303 Grades	1.0	1	2.00	
	1.4	5	3.20	0.84
	2.0	3	2.33	1.15
	2.1	3	2.33	1.15
	2.9	3	3.00	1.00
	3.0	11	2.91	1.38
	3.7	11	3.18	0.75
	4.0	10	3.10	1.20
	4.5	17	4.00	0.00
	4.8	8	4.00	0.00
	5.0	10	4.00	0.00
Other Biology Hours Taken	1.0	2	4.00	2.83
	1.4	8	4.50	1.85
	2.0	5	2.60	0.55
	2.1	2	4.00	2.83
	2.9	3	2.00	0.00
	3.0	6	2.67	1.63
	3.7	4	2.00	0.00
	4.0	5	2.80	1.30
	4.5	2	2.00	0.00
	4.8	3	2.00	0.00
	5.0	2	2.50	0.71
GPAs in Other Biology Classes	1.0	2	4.00	0.00
	1.4	8	3.31	0.88
	2.0	5	3.00	0.00
	2.1	2	2.58	0.82
	2.9	3	3.00	1.00
	3.0	6	3.83	0.41
	3.7	4	3.00	0.00
	4.0	5	3.80	0.45
	4.5	2	3.00	1.41
	4.8	3	3.67	0.58
	5.0	2	3.50	0.71

TABLE A.27

Biology 1996: Descriptive Statistics of the Biology Outcome Measures

	Mean AP Score Within Cluster	*Frequency*	*Mean*	*Standard Deviation*
BIO 303 Grades	1.0	4	2.25	0.50
	1.4	7	2.29	0.76
	1.9	14	2.36	1.01
	2.2	22	2.32	1.39
	2.8	22	3.00	0.93
	3.2	21	3.10	0.94
	3.5	17	3.71	0.59
	4.0	30	3.37	0.93
	4.5	23	3.35	0.98
	4.8	47	4.00	0.00
	5.0	24	4.00	0.00
Other Biology Hours Taken	1.0	10	6.00	2.36
	1.4	16	4.63	2.16
	1.9	15	4.80	3.03
	2.2	19	3.37	1.67
	2.8	16	3.75	2.35
	3.2	15	3.33	2.35
	3.5	7	4.43	5.13
	4.0	15	3.40	2.13
	4.5	11	3.82	2.89
	4.8	20	4.05	3.62
	5.0	10	3.40	2.41
GPAs in Other Biology Classes	1.0	10	2.32	0.61
	1.4	16	2.53	1.16
	1.9	15	2.99	0.87
	2.2	19	2.75	1.13
	2.8	16	3.08	0.67
	3.2	15	2.54	0.92
	3.5	7	3.56	0.78
	4.0	15	3.54	0.62
	4.5	11	3.23	0.63
	4.8	20	3.64	0.69
	5.0	10	3.55	0.60

BIO 303 Grades ($F = 12.97$, $p < .001$)	5.0 > 1.0 – 3.2 4.8 > 1.0 – 3.2, 4.0 4.0, 4.5 > 1.9, 2.2 3.5 > 1.0 – 2.2
Other GPAs ($F = 4.34$, $p < .001$):	5.0 > 1.9 4.8 > 1.0, 1.4, 2.2, 3.2 4.0 > 1.0, 1.4, 3.2

TABLE A.28

Biology 1997: Descriptive Statistics of the Biology Outcome Measures

	Mean AP Score Within Cluster	*Frequency*	*Mean*	*Standard Deviation*
BIO 303 Grades	1.0	2	2.00	0.00
	1.4	9	2.44	1.13
	2.0	12	2.17	1.40
	2.2	9	3.00	0.87
	2.9	20	2.35	1.50
	3.1	23	2.48	0.90
	3.8	25	3.48	1.19
	3.9	23	3.13	0.69
	4.3	38	3.68	0.81
	4.8	42	4.00	0.00
	5.0	32	4.00	0.00
Other Biology Hours Taken	1.0	11	5.00	1.90
	1.4	20	4.25	2.05
	2.0	12	5.17	3.51
	2.2	23	4.74	2.83
	2.9	14	5.71	4.65
	3.1	22	4.14	2.68
	3.8	16	5.50	3.63
	3.9	21	5.86	3.77
	4.3	17	4.24	2.84
	4.8	17	4.88	2.91
	5.0	18	5.06	3.78
GPAs in Other Biology Classes	1.0	11	2.18	1.40
	1.4	20	2.68	1.22
	2.0	12	2.94	0.88
	2.2	23	2.67	1.25
	2.9	14	3.03	0.87
	3.1	22	3.00	0.92
	3.8	16	2.96	1.05
	3.9	21	3.03	0.75
	4.3	17	3.38	1.09
	4.8	17	3.25	1.02
	5.0	18	3.62	0.76

BIO 303 Grades ($F = 13.81$, $p < .001$): 4.8, 5.0 > 1.4, 2.0, 2.9, 3.1, 3.9
4.3 > 1.4, 2.0, 2.9, 3.1
3.8 > 2.0, 3.1

TABLE A.29

Biology 1998: Descriptive Statistics of the Biology Outcome Measures

	Mean AP Score Within Cluster	*Frequency*	*Mean*	*Standard Deviation*
BIO 303 Grades	1.0	7	2.29	1.11
	1.3	12	2.75	1.14
	1.8	10	2.80	0.63
	2.1	17	2.94	0.83
	2.6	13	2.46	0.88
	3.1	20	3.05	1.28
	3.5	18	2.94	1.00
	4.2	32	3.06	1.32
	4.6	27	3.74	0.81
	5.0(a)	37	4.00	0.00
	5.0(b)	25	4.00	0.00
Other Biology Hours Taken	1.0	11	4.36	1.43
	1.3	24	4.75	2.31
	1.8	21	4.05	1.75
	2.1	16	4.75	1.84
	2.6	13	3.23	1.17
	3.1	24	3.79	3.18
	3.5	14	3.79	2.08
	4.2	19	5.05	2.57
	4.6	14	3.64	2.59
	5.0(a)	12	4.08	2.02
	5.0(b)	13	4.23	2.71
GPAs in Other Biology Classes	1.0	11	2.59	1.07
	1.3	24	3.05	0.80
	1.8	21	3.19	0.61
	2.1	16	2.87	1.10
	2.6	13	2.37	0.98
	3.1	24	2.83	1.13
	3.5	14	2.94	1.22
	4.2	19	3.36	0.79
	4.6	14	3.26	0.93
	5.0(a)	12	3.60	0.58
	5.0(b)	13	3.41	0.50

BIO 303 Grades ($F = 8.26$, $p < .001$): 5.0(a), 5.0(b) > 1.0 – 4.2
4.6 > 1.0, 2.61

TABLE A.30

Biology 1999: Descriptive Statistics of the Biology Outcome Measures

	Mean AP Score Within Cluster	*Frequency*	*Mean*	*Standard Deviation*
BIO 303 Grades	1.0	1	0.00	
	1.7	7	2.57	0.53
	2.3	8	2.63	1.30
	2.5	8	1.75	1.28
	3.2	14	2.36	1.55
	3.6	8	3.50	0.76
	3.7	9	3.44	0.73
	4.2	22	3.68	0.57
	4.7	15	3.87	0.52
	5.0(a)	27	4.00	0.00
	5.0(b)	12	4.00	0.00
Other Biology Hours Taken	1.0	17	4.00	2.15
	1.7	26	4.31	1.38
	2.3	15	3.27	1.33
	2.5	13	3.85	1.72
	3.2	11	4.09	1.70
	3.6	8	3.25	0.89
	3.7	9	3.56	1.67
	4.2	12	3.08	1.31
	4.7	7	3.57	1.40
	5.0(a)	11	4.18	1.78
	5.0(b)	6	4.17	2.32
GPAs in Other Biology Classes	1.0	17	2.65	1.07
	1.7	26	2.63	0.87
	2.3	15	3.00	1.02
	2.5	13	2.55	1.08
	3.2	11	2.55	0.98
	3.6	8	2.72	1.30
	3.7	9	2.85	1.06
	4.2	12	2.78	1.41
	4.7	7	3.26	0.55
	5.0(a)	11	3.58	0.67
	5.0(b)	6	4.00	0.00