



To: Charles Curtis, Chair, CEP
From: David Crowne, Alan Houston, Sandy Lakoff, Paul Libby, Michael Schudson
Re: Freshman Writing Evaluation
2 May 2006

This committee is composed of five faculty members who agreed to evaluate, individually and independently, 384 freshman essays from the six college core curricula or writing programs. Mark Appelbaum's office collected the essays, made them available to us, and provided us with a grading rubric. (See Appendix A). Each reviewer read sixteen student portfolios, each comprising four papers, from each of the six colleges. Each portfolio, in turn, was read by two independent reviewers. The committee met in July 2005 to distribute portfolios, and again in February 2006 to discuss the results of the study.

The data that accompany this report were compiled and analyzed by Mark Appelbaum and his staff. CEP should of course evaluate the data on its own, and consult with Mark as the best informed guide to the selection of the data and the analysis. However, we have a few thoughts of our own that may be useful to you.

Results

Three general results of this study seem to us the most important:

1. *The overall quality of freshman writing is satisfactory.* Each reviewer assigned each essay a Subject A score (6 – 1). Each portfolio was also assigned a global score using a traditional letter scale (A – F). The mean Subject A scores for the entire pool of essays ranged from 3.99 on assignment A1 to 4.14 on assignment B2. (See Tables 9 – 12). According to the rubric, a paper with a score of 4 is “satisfactory, sometimes marginally so.” The mean global score for the entire pool of essays was 2.81, corresponding to a letter grade of “B-.” (See Table 13).
2. *The quality of student writing varied across the six colleges.* The global score had a mean of 3.09 for Revelle, 2.93 for ERC, 2.88 for Marshall, 2.87 for Muir, 2.66 for Sixth, and 2.56 for Warren. (See Table 19). These scores correspond to letter grades ranging from “B” (Revelle) to “C+/B-” (Warren). These differences are statistically significant.
3. *The quality of student writing did not improve significantly.* The mean improvement score for the entire pool of essays was 2.22. (See Table 14). According to the rubric, a portfolio of papers with a score of 2 shows “no improvement, writing is essentially unchanged.” There were no statistically significant differences among the colleges in improvement scores.

Analysis

Concerning variations in outcome among the colleges:

We do not know what differences among the colleges are responsible for the varying quality of student writing. We considered the possibility that these variations reflected differences among the students entering each of the colleges. Analysis by Mark Appelbaum, however, has ruled out the most likely such differences (parents' education, family income, high school GPA, SAT scores, cumulative spring quarter GPA at UCSD). The colleges did not differ significantly in any of these variables. Colleges differ greatly by gender composition -- ERC is 28% male, while Warren is 68% male -- but with only small differences among high school GPA or SAT verbal scores, it is not clear how gender composition would explain the results.

Other variables that were not measured may also have been significant. Perhaps there is an uneven distribution across colleges of students whose first language is not English. Surely there is an uneven distribution across colleges of students apt by area of academic interest to have weaker skills in writing or, for that matter, apt to have more demanding courses to work on apart from their freshman writing course.

The committee operated within a "black box," evaluating essays without knowledge of or reference to the specific methods used to improve student writing. The only aspect of instruction we could observe from our perch is how the "prompts" or essay assignments at each college were written. Our own sense is that comparing these prompts demonstrated greater clarity, simplicity and focus at Revelle. As the prompts grew more ambiguous and vague, or more ambitious and complex, students' writing weakened. There may be good pedagogical reasons for such prompts, but they appear to exact a cost. This is our own subjective sense; it is neither supported nor refuted by the data.

Concerning the improvement finding:

The mean improvement score was 2.22. The mean of individual reviewers differed significantly, from a low of 1.87 to a high of 2.66. (See Table 7). However, there were no statistically significant differences among the colleges; mean scores ranged from 2.15 (Muir) to 2.32 (Revelle). (See Tables 20, A1 and A2). Moreover, there were no statistically significant relationships between improvement scores and the grades students received in their writing courses. (See Table A3).

We do not know why we detected so little improvement. It is possible that improvement did occur, but was invisible as a consequence of the way the portfolios were assembled. In almost all cases, the essays we read were "final drafts," and had already gone through a process of revision. We did not read "raw" or unedited student work. The greatest changes in student writing may occur during the transformation from first to final draft.

Nonetheless, we would hope to see improvement in final drafts from the first essay, A1, to the final essay, B2. It is possible that what students have learned over twelve years of schooling is not easily altered in two years of intensive writing instruction in the thirteenth year. It is also possible, of course, that the colleges do not do a good job of teaching freshman writing. The data do not offer any guidance among these possibilities.

Recommendations

It would be valuable to know if writing instruction at other universities has shown greater signs of student progress. Unfortunately, we are not aware of any studies comparable to this one.

Members of the committee are divided on the advisability of repeating this study. Some members feel that this report, in combination with previous evaluations of UCSD writing instruction, is sufficient. Others feel that findings regarding variations among the colleges and lack of student improvement merit further study. We take no position on the merits of conducting a similar study of students selected on a different basis (e.g., English-language learners).

Should further study be called for, better training might reduce variation among reviewers. At our initial meeting we were provided brief descriptions of each point on the numerical Subject A grading scale. These descriptions seemed sensible and reflected our intuitive notions of better and worse writing. As tables 2 – 7 indicate, however, we differed – sometimes significantly – in the application of the scale. It is possible that additional training might bring reviewers into greater consensus. We would caution, however, that much of the impetus for evaluating the writing programs came from the subjective sense of the faculty that “our students do not write well.” Faculty come to these conclusions on their own, without any more training or experience than the members of our committee. If reviewers were too well trained – if they followed more closely a particular model of quality writing that did not map onto the subjective sense of the faculty at large – then the results might not be as compelling to members of the Senate as the procedure used in this exercise.

What impact might greater training have had on our findings? If we had graded more uniformly than we did, then the overall global score of 2.81 might have been a little higher or a little lower. Differences among the colleges might have been a little larger or a little smaller, though it is difficult to imagine that they would have been erased, or that their direction would have been altered. We do not believe that the improvement score would be significantly altered. We do, however, leave open the possibility that improvement scores would have looked different had we read “raw” essays and not final drafts.

Freshman Writing Analysis – Supplement 1

At the request of the Committee a supplemental analysis designed to isolate possible factors that might identify students with very high or very low improvement scores was undertaken. To this end students with mean improvement scores below 2.0 (approximately the lowest 10% of students in the sample) were coded with a special analysis variable LowImprovement = 1 and students with mean improvement scores above 2.85 (approximately the upper 10% of students) were coded with a special analysis variable HighImprovement = 1. Analyses of those students with low improvement scores (n=22 for students with mean improvement scores below 2.0) are presented in Table A1; analyses of those high improvement scores (n=26 for students with mean improvement scores above 2.85) are presented in Table A2; and an analysis of the relationship between course grades and improvement groups is presented in Table A3.

Low Improvement differences (Table A1): No statistically significant differences (but recall the sample sizes are very small) were detected in terms of differential rates of low improvement scores among the six programs of instruction. There were, likewise, no statistically significant differences between the low improvement group and all other students in the sample on any of the variables: Father Ed, Mother Ed, Income, High School GPA, SAT-V, SAT-M, and Spring Cumulative GPA.

High Improvement differences (Table A2): No statistically significant differences (but recall the sample sizes are very small) were detected in terms of differential rates of high improvement scores among the six programs of instruction. There were, likewise, no statistically significant differences between the high improvement group and all other students in the sample on any of the variables: Father Ed, Mother Ed, Income, High School GPA, SAT-V, and SAT-M. There was, however, a statistically significant difference between the high improvement students (gpa = 3.28) and all other students in the sample (gpa = 3.04) on Spring Cumulative GPA which might lead to a speculation that students who show the greatest writing improvement are working somewhat harder on their academic activities than are students in general. {The gpa difference is greater than can be accounted for by just the difference in student grades in the writing courses.}

Relation between Improvement groups and grades in courses (Table A3): No statistically significant relationships were observed between membership in either the high improvement or low improvement group and grade in the writing courses.

Table A1: Students with Mean Improvement Scores below 2.0

By College:

Contingency Table

LowImprovement By College

Count	E	M	R	S	T	W	
Total %							
Col %							
Row %							
0	37 15.42 92.50 16.97	37 15.42 92.50 16.97	37 15.42 92.50 16.97	37 15.42 92.50 16.97	34 14.17 85.00 15.60	36 15.00 90.00 16.51	218 90.83
1	3 1.25 7.50 13.64	3 1.25 7.50 13.64	3 1.25 7.50 13.64	3 1.25 7.50 13.64	6 2.50 15.00 27.27	4 1.67 10.00 18.18	22 9.17
	40 16.67	40 16.67	40 16.67	40 16.67	40 16.67	40 16.67	240

Test	ChiSquare	Prob>ChiSq
Likelihood Ratio	1.995	0.8499
Pearson	2.202	0.8206

Oneway Analysis of FATHER ED By LowImprovement

Oneway Anova

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
LowImprovement	1	1.4261	1.42612	0.2675	0.6056
Error	203	1082.2519	5.33129		
C. Total	204	1083.6780			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
0	187	5.85027	0.16885	5.5173	6.1832
1	18	5.55556	0.54423	4.4825	6.6286

Std Error uses a pooled estimate of error variance

Oneway Analysis of MOTHER ED By LowImprovement

Oneway Anova

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
LowImprovement	1	12.7522	12.7522	2.5819	0.1096
Error	211	1042.1585	4.9391		
C. Total	212	1054.9108			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
0	193	5.53886	0.15997	5.2235	5.8542
1	20	4.70000	0.49695	3.7204	5.6796

Std Error uses a pooled estimate of error variance

Oneway Analysis of INCOME By LowImprovement**Oneway Anova****Analysis of Variance**

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
LowImprovement	1	89255875	89255875	0.0196	0.8888
Error	177	8.06408e11	4.55598e9		
C. Total	178	8.06497e11			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
0	162	77549.3	5303	67084	88015
1	17	75140.7	16371	42834	107448

Std Error uses a pooled estimate of error variance

Oneway Analysis of HS_GPA By LowImprovement**Oneway Anova****Analysis of Variance**

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
LowImprovement	1	0.018738	0.018738	0.1433	0.7054
Error	219	28.641009	0.130781		
C. Total	220	28.659747			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
0	200	3.97140	0.02557	3.9210	4.0218
1	21	3.94000	0.07892	3.7845	4.0955

Std Error uses a pooled estimate of error variance

Oneway Analysis of SATVOFL By LowImprovement**Oneway Anova****Analysis of Variance**

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
LowImprovement	1	8459.9	8459.85	0.9906	0.3207
Error	222	1895961.6	8540.37		
C. Total	223	1904421.4			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
0	203	591.084	6.486	578.30	603.87
1	21	570.000	20.166	530.26	609.74

Std Error uses a pooled estimate of error variance

Oneway Analysis of SATMOFL By LowImprovement**Oneway Anova****Analysis of Variance**

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
LowImprovement	1	1786.2	1786.23	0.1823	0.6698
Error	222	2175026.3	9797.42		
C. Total	223	2176812.5			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
0	203	637.783	6.947	624.09	651.47
1	21	628.095	21.600	585.53	670.66

Std Error uses a pooled estimate of error variance

Oneway Analysis of SP05 CUMGPA By LowImprovement

Oneway Anova

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
LowImprovement	1	0.041149	0.041149	0.1911	0.6625
Error	222	47.810386	0.215362		
C. Total	223	47.851535			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
0	203	3.06554	0.03257	3.0014	3.1297
1	21	3.11204	0.10127	2.9125	3.3116

Std Error uses a pooled estimate of error variance

Table 2A: Highest 10% of Students on Mean Improvement Scores (mean improvement greater than 2.85)

HilImprovement By College

Count	E	M	R	S	T	W	
Total %							
Col %							
Row %							
0	37	38	34	35	36	34	214
	15.42	15.83	14.17	14.58	15.00	14.17	89.17
	92.50	95.00	85.00	87.50	90.00	85.00	
	17.29	17.76	15.89	16.36	16.82	15.89	
1	3	2	6	5	4	6	26
	1.25	0.83	2.50	2.08	1.67	2.50	10.83
	7.50	5.00	15.00	12.50	10.00	15.00	
	11.54	7.69	23.08	19.23	15.38	23.08	
	40	40	40	40	40	40	240
	16.67	16.67	16.67	16.67	16.67	16.67	

Test ChiSquare Prob>ChiSq
 Likelihood Ratio 3.674 0.5972
 Pearson 3.451 0.6309
 Warning: 20% of cells have expected count less than 5, ChiSquare suspect

Oneway Analysis of FATHER ED By HilImprovement

Oneway Anova

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
HilImprovement	1	19.2839	19.2839	3.6778	0.0565
Error	203	1064.3941	5.2433		
C. Total	204	1083.6780			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
0	181	5.71271	0.17020	5.3771	6.0483
1	24	6.66667	0.46741	5.7451	7.5883

Std Error uses a pooled estimate of error variance

Oneway Analysis of MOTHER ED By HilImprovement

Oneway Anova

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
HilImprovement	1	9.5253	9.52527	1.9226	0.1670
Error	211	1045.3855	4.95443		
C. Total	212	1054.9108			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
0	188	5.38298	0.16234	5.0630	5.7030
1	25	6.04000	0.44517	5.1624	6.9176

Std Error uses a pooled estimate of error variance

Oneway Analysis of INCOME By HilImprovement**Oneway Anova****Analysis of Variance**

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
HilImprovement	1	165082750	165082750	0.0362	0.8492
Error	177	8.06332e11	4.55555e9		
C. Total	178	8.06497e11			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
0	158	77670.7	5370	67074	88267
1	21	74686.4	14729	45620	103753

Std Error uses a pooled estimate of error variance

Oneway Analysis of HS_GPA By HilImprovement**Oneway Anova****Analysis of Variance**

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
HilImprovement	1	0.351286	0.351286	2.7176	0.1007
Error	219	28.308460	0.129262		
C. Total	220	28.659747			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
0	195	3.98297	0.02575	3.9322	4.0337
1	26	3.85923	0.07051	3.7203	3.9982

Std Error uses a pooled estimate of error variance

Oneway Analysis of SATVOFL By HilImprovement**Oneway Anova****Analysis of Variance**

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
HilImprovement	1	10584.8	10584.8	1.2408	0.2665
Error	222	1893836.7	8530.8		
C. Total	223	1904421.4			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
0	198	586.616	6.564	573.68	599.55
1	26	608.077	18.114	572.38	643.77

Std Error uses a pooled estimate of error variance

Oneway Analysis of SATMOFL By HilImprovement**Oneway Anova****Analysis of Variance**

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
HilImprovement	1	27241.9	27241.9	2.8134	0.0949
Error	222	2149570.6	9682.8		
C. Total	223	2176812.5			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
0	198	632.879	6.993	619.10	646.66
1	26	667.308	19.298	629.28	705.34

Std Error uses a pooled estimate of error variance

Oneway Analysis of SP05 CUMGPA By HilImprovement

Oneway Anova

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
HilImprovement	1	1.245105	1.24511	5.9308	0.0157
Error	222	46.606430	0.20994		
C. Total	223	47.851535			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
0	198	3.04289	0.03256	2.9787	3.1071
1	26	3.27565	0.08986	3.0986	3.4527

Std Error uses a pooled estimate of error variance

Table 3: Course Grades as a function of High or Low Improvement Scores

Oneway Analysis of Course 1 Grade By HilImprovement

Oneway Anova

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
HilImprovement	1	0.701701	0.701701	2.1110	0.1476
Error	238	79.111633	0.332402		
C. Total	239	79.813334			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
0	214	2.91449	0.03941	2.8368	2.9921
1	26	3.08846	0.11307	2.8657	3.3112

Std Error uses a pooled estimate of error variance

Oneway Analysis of Course 2 Grade By HilImprovement

Oneway Anova

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
HilImprovement	1	0.809016	0.809016	2.2806	0.1323
Error	236	83.717413	0.354735		
C. Total	237	84.526429			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
0	212	2.94387	0.04091	2.8633	3.0245
1	26	3.13077	0.11681	2.9007	3.3609

Std Error uses a pooled estimate of error variance

Oneway Analysis of Course 1 Grade By LowImprovement

Oneway Anova

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
LowImprovement	1	0.064276	0.064276	0.1918	0.6618
Error	238	79.749058	0.335080		
C. Total	239	79.813334			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
0	218	2.93853	0.03921	2.8613	3.0158
1	22	2.88182	0.12341	2.6387	3.1249

Std Error uses a pooled estimate of error variance

Oneway Analysis of Course 2 Grade By LowImprovement

Oneway Anova

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
LowImprovement	1	0.033209	0.033209	0.0928	0.7610
Error	236	84.493220	0.358022		
C. Total	237	84.526429			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
0	216	2.96806	0.04071	2.8878	3.0483
1	22	2.92727	0.12757	2.6760	3.1786

Std Error uses a pooled estimate of error variance

Assessment of Freshman Writing
 A Report on the 2004/2005 Sample

In accordance with the proposal initiated by the Committee on Educational Policy designed to assess the state of freshman writing at UCSD, a sample of the writing products of students enrolled in the six core curricula (during those quarters specified as the “writing quarters”) or writing programs were collected. The sample was collected by having each student in the writing program, during the winter and spring quarters of the 2004/2005 academic year, deposit their writing samples into an electronic library created for this purpose. Students were instructed to deposit four writing samples over the course of the two quarters – an early writing assignment from the winter quarter course, the final writing sample of the winter quarter course, an early writing assignment for the spring quarter course, and the final writing assignment from the spring quarter course. The rates of deposit are given in the Table 1. With the exception of Warren Writing, the level of deposit was considered sufficient for the purpose of this study. {There is no indication that the differences between those who did elect to deposit their writing samples and those who did not was sufficiently great to bias the results of the study. There were no significant differences on SAT V and course grades between those who did and did not deposit for the five colleges with sufficient data to draw a conclusion. }

Table 1: Deposit rates

	Enroll		1st		2nd	
CAT 2	449	242	53.9%	242	53.9%	
DOC 2	751	690	91.9%	636	84.7%	
HUM 1	693	598	86.3%	556	80.2%	
MCWP 40	233	182	78.1%	177	76.0%	
MMW 2	745	645	86.6%	607	81.5%	
WCWP 10A	179	100	55.9%	94	52.5%	

	Enroll		1st		2nd	
CAT3	445	429	96.4%	391	87.9%	
DOC3	776	716	92.3%	685	88.3%	
HUM2	644	525	81.5%	466	72.4%	
MCWP50	306	197	64.4%	198	64.7%	
MMW3	781	650	83.2%	579	74.1%	
WCWP10B	334	140	41.9%	72	21.6%	

For the five colleges with sufficiently rich deposit rates (better than 50% for every assignment) simple random samples of portfolios were drawn. Initially 60 portfolios were drawn from each college and the first 40 that met the criterion of having a complete set of four writing assignments were retained. Since there were not a sufficient number of electronically deposited writing samples for Warren, we took advantage of the fact that

Warren Writing had retained the writing products of all of its students. Therefore, in order to obtain an adequate sample of portfolios, we generated a random list of 60 students who were registered (at third week) for both Warren 10A and 10B. The staff at Warren provided copies of the writing for the first 40 student who met the criterion of having all four writing elements.

Ratings of the writing samples

Five raters were recruited by the Academic Senate's Committee on Committees to read and assess each writing portfolio. These individuals are all Academic Senate Faculty with no direct association with any of the six core curriculum/writing programs. During the summer of 2005 these faculty members were commissioned to read a subset of the portfolios and to provide six ratings for each portfolio: four ratings of the individual writing exemplars graded on the "Subject A" Scale, a global rating of the four writing pieces taken together on an A-F scale, and a global rating of the amount of improvement seen in the writing. Copies of the instructions received by the raters and the formal definitions for the "Subject A" Scale are given in Appendix A.

Assignment of portfolios to reviewers

Portfolios were randomly assigned to the five reviewers in such a manner that (1) every portfolio would be read independently by two reviewers, (2) every reviewer would review the same number of portfolios from each program, and (3) each program would be reviewed equally often by each reviewer. This stratification scheme was employed to assure that any differences between programs that might be observed were not due to rater differences alone. The actual randomization of portfolios to reviewers was conducted early in the summer of 2005 (after all actual student grades on this work had been assigned by the instructor and had been officially recorded – i.e. to protect the student from any "risks" that participation in this project might have engendered). Each reviewer (referred to later in this report by a color code which corresponds to the color of the folders in which their reading samples were kept) received 96 portfolios (16 from each of the six programs) to review and rate. Since each portfolio contained four writing samples (except those from one program which has only a single major "term paper" writing project in its second quarter) each reviewer was asked to read 384 writing items and to provide 576 individual assessments.

Rater Differences

Since raters only met one time (and never as a complete group) to discuss the rating process and were largely instructed to use the written materials inform their judgment process, it is important to examine for rater differences. Tables 2-7 provide rater specific information of each of the six ratings that the raters were asked to provide. For each of the six ratings, there were highly significant differences in the ways that the raters assigned the ratings and in all cases the level of statistical significance was at the .0001 level of significance. Rater "B" consistently assigned the lowest ratings of any of the reviewers and was often statistically different from the other four raters. Similarly Rater

“Y” assigned consistently higher ratings for the ratings of the individual writing samples on using the “Subject A” criteria but was near the middle of the group on the Global and Improvement ratings.

Reliability of ratings

(To be provided)

Overall Ratings

The full sample results of the reading and assessment process are given in Tables 9 – 14. A simple overview of these results can be had by considering the 75th, 50th, and 25th percentile scores of the ratings on each of the six ratings. These percentiles are computed on the mean score of the two ratings for each portfolio on each rating.

Scale	25 th	50 th	75 th
Subject A Rating – A1	3.5	4.0	4.5
Subject A Rating – A2	3.5	4.0	4.6
Subject A Rating - B1	3.5	4.0	4.7
Subject A Rating - B2	3.5	4.1	4.9
Global Rating	2.35	2.85	3.35
Improvement Rating	2.0	2.0	2.5

Program Differences

The results of analyses examining the differences between the six college level programs are given in Tables 15-20. On four of the six rating scales, Mean of A1, Mean of B1, Mean of B2, and Mean of Global, there were statistically significant differences between the writing programs. Ratings given to student writing samples from Revelle were consistently the highest (on average) while those given to the writing samples from Sixth and Warren were consistently the lowest (on average). There were, however, no statistically significant differences among the colleges on the Mean of A2 and on Mean Improvement.

The detection of statistically significant differences between the programs immediately begs the question – do the students in the samples of the several programs differ before the fact on factors known (or thought) to be related to student academic achievement. To this end analyses were conducted to determine if there were sample differences on the following variables:

Father Educational Level

Mother Educational Level
Reported Family Income
High School GPA
SAT – Verbal
SAT – Math
Cumulative Spring Quarter GPA at UCSD
Gender

The detailed results of these inquiries are provided in Appendix B. Suffice it to say that, with the exception of Gender, there were no Program differences on these variables. There also appeared to be no indication of Program differences in the ESL/Subject A rates among the programs.

Correlates of Ratings

In Table 21 the correlations between the Global Ratings and the Improvement Ratings and the achieved course grades are presented.

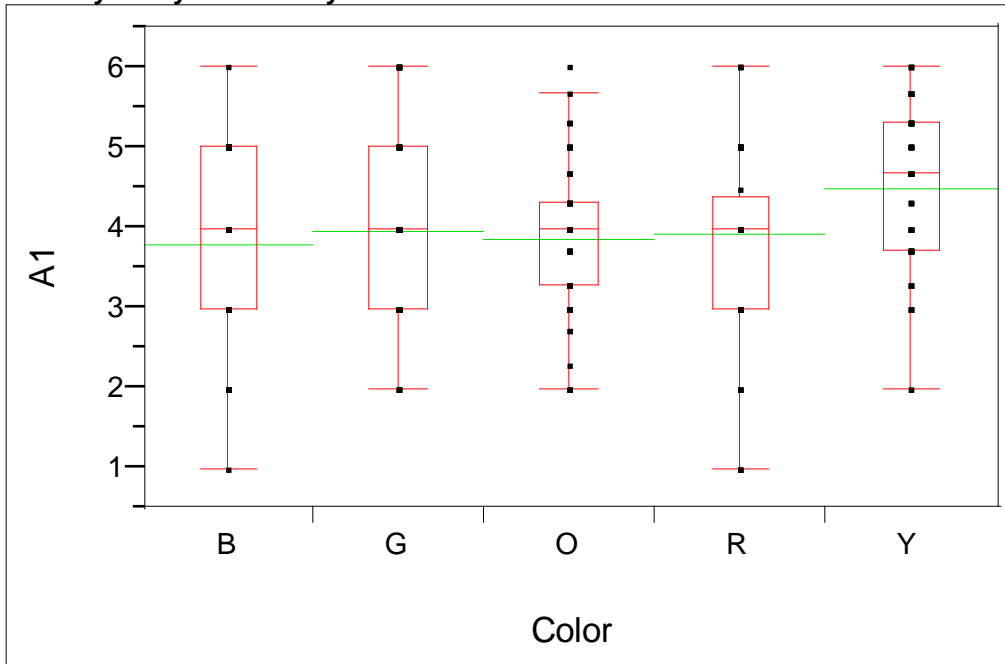
In Table 22 the correlations between the Global Ratings and several entering characteristics of the students in the sample (Father Education, Mother Education, Reported Family Income, HS GPA, SAT-V, and SAT-M) are presented.

In Table 23 the correlations between the Improvement Score and the same entering characteristics are presented.

In Table 24 the correlations between Achieved Grades and the same entering characteristics are presented.

Analysis Tables

Table 2
Oneway Analysis of A1 By Color



Oneway Anova

Analysis of Variance

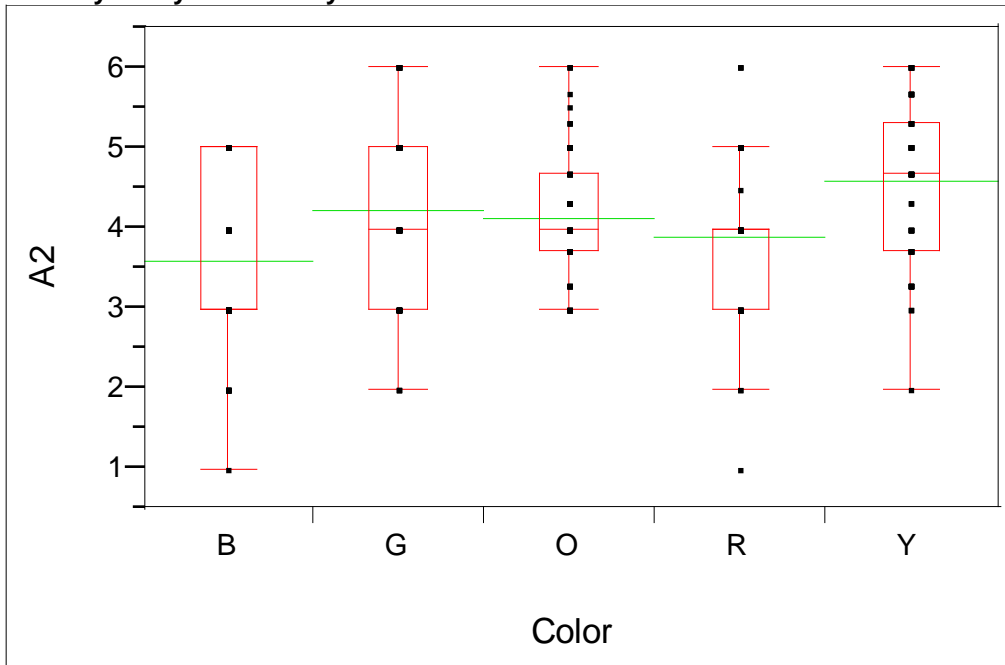
Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Color	4	29.80024	7.45006	7.5384	<.0001
Error	474	468.44339	0.98828		
C. Total	478	498.24363			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
B	95	3.76842	0.10199	3.5680	3.9688
G	96	3.93750	0.10146	3.7381	4.1369
O	96	3.86250	0.10146	3.6631	4.0619
R	96	3.92188	0.10146	3.7225	4.1212
Y	96	4.47812	0.10146	4.2788	4.6775

Std Error uses a pooled estimate of error variance

Table 3
Oneway Analysis of A2 By Color



Oneway Anova

Analysis of Variance

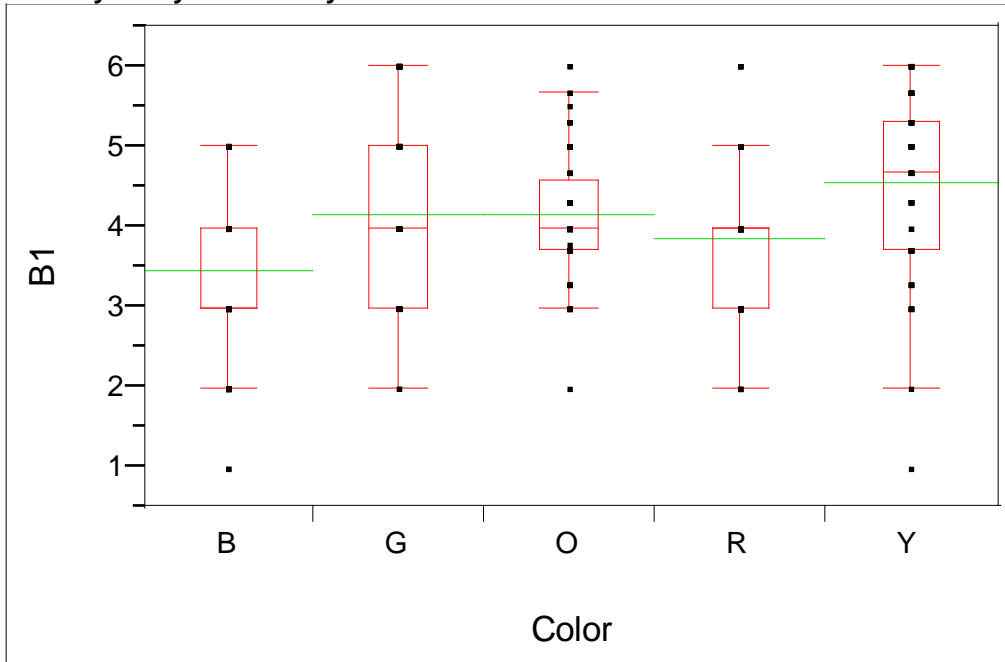
Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Color	4	52.03473	13.0087	14.4550	<.0001
Error	474	426.57458	0.8999		
C. Total	478	478.60931			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
B	95	3.58947	0.09733	3.3982	3.7807
G	96	4.21875	0.09682	4.0285	4.4090
O	96	4.10938	0.09682	3.9191	4.2996
R	96	3.88021	0.09682	3.6900	4.0705
Y	96	4.57396	0.09682	4.3837	4.7642

Std Error uses a pooled estimate of error variance

Table 4
Oneway Analysis of B1 By Color



Oneway Anova

Analysis of Variance

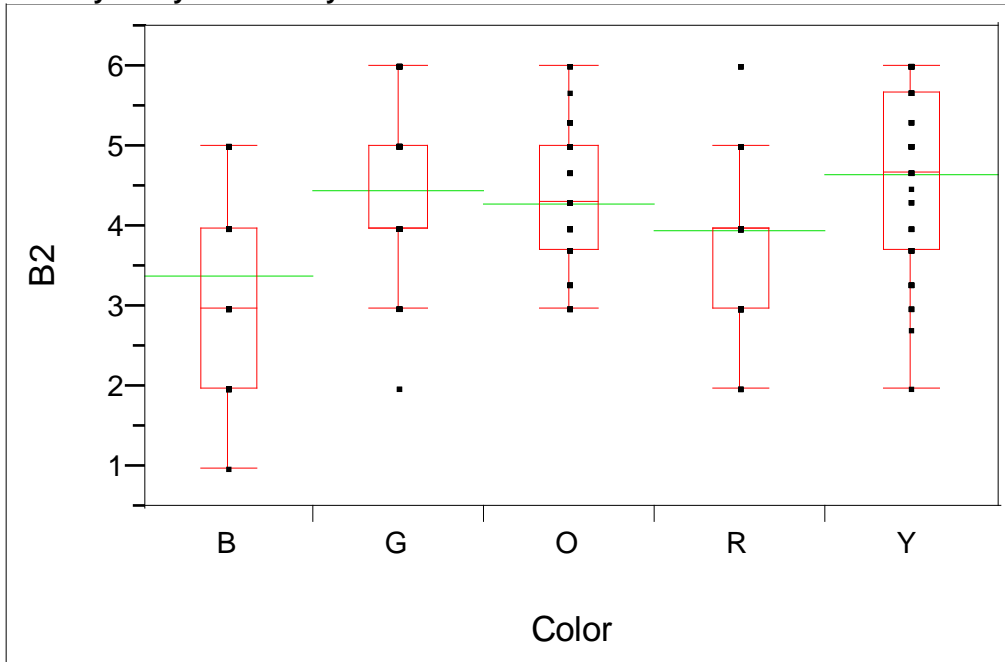
Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Color	4	65.33333	16.3333	16.5865	<.0001
Error	474	466.76413	0.9847		
C. Total	478	532.09746			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
B	95	3.44211	0.10181	3.2420	3.6422
G	96	4.15625	0.10128	3.9572	4.3553
O	96	4.15385	0.10128	3.9548	4.3529
R	96	3.84375	0.10128	3.6447	4.0428
Y	96	4.55208	0.10128	4.3531	4.7511

Std Error uses a pooled estimate of error variance

Table 5
Oneway Analysis of B2 By Color



Oneway Anova

Analysis of Variance

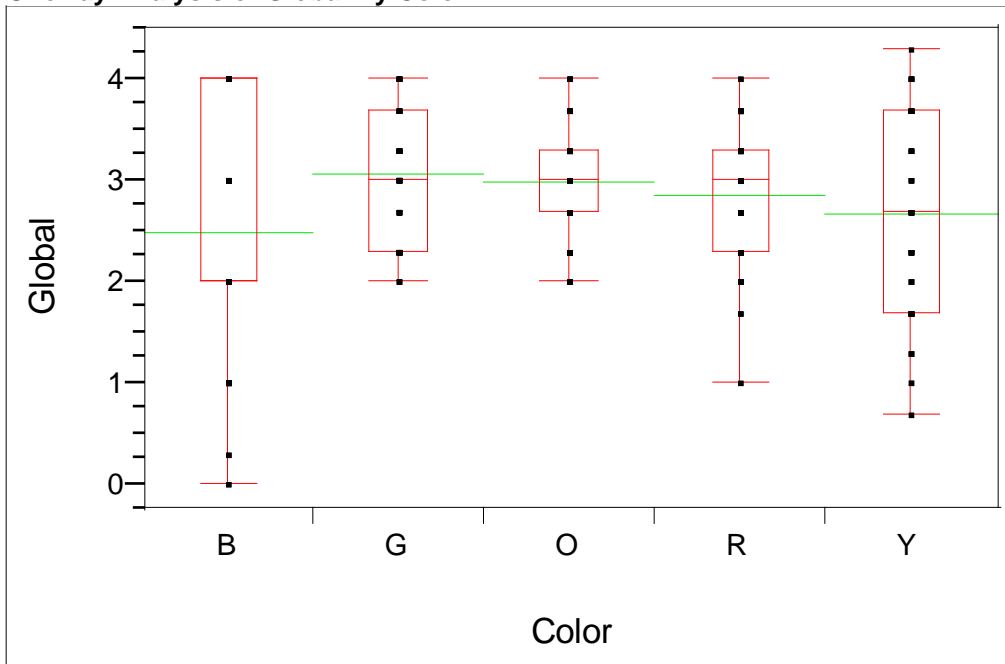
Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Color	4	91.78997	22.9475	23.1042	<.0001
Error	474	470.78551	0.9932		
C. Total	478	562.57549			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
B	95	3.38947	0.10225	3.1886	3.5904
G	96	4.44792	0.10172	4.2480	4.6478
O	96	4.27604	0.10172	4.0762	4.4759
R	96	3.93750	0.10172	3.7376	4.1374
Y	96	4.63437	0.10172	4.4345	4.8342

Std Error uses a pooled estimate of error variance

Table 6
Oneway Analysis of Global By Color



Oneway Anova

Analysis of Variance

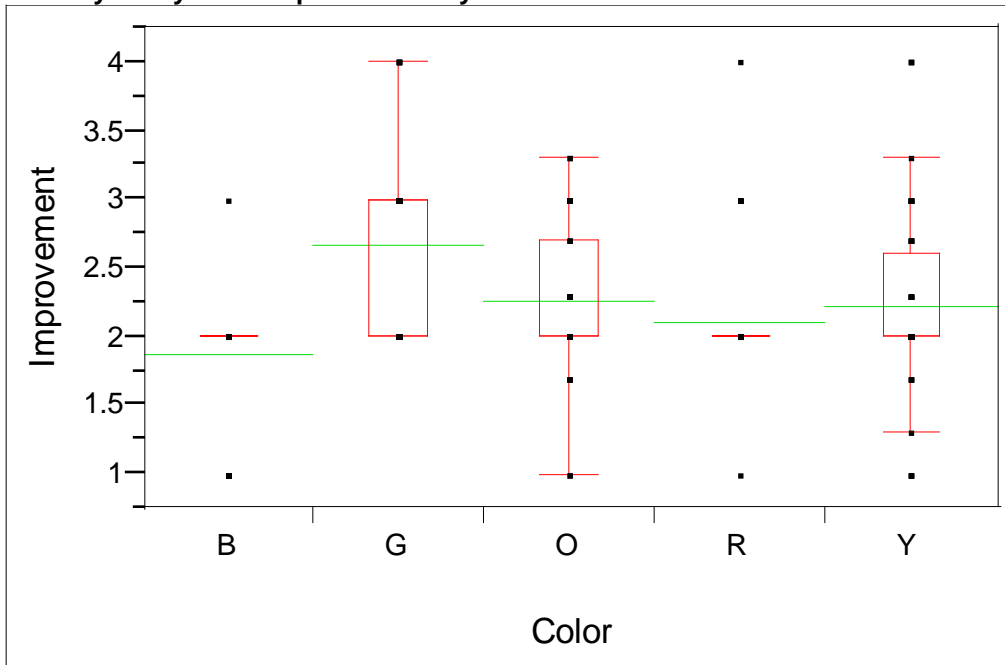
Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Color	4	21.85520	5.46380	7.9145	<.0001
Error	474	327.22697	0.69035		
C. Total	478	349.08218			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
B	95	2.47263	0.08525	2.3051	2.6401
G	96	3.06042	0.08480	2.8938	3.2270
O	96	2.99062	0.08480	2.8240	3.1573
R	96	2.85521	0.08480	2.6886	3.0218
Y	96	2.67708	0.08480	2.5105	2.8437

Std Error uses a pooled estimate of error variance

Table 7
Oneway Analysis of Improvement By Color



Oneway Anova

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Color	4	31.31335	7.82834	31.7355	<.0001
Error	474	116.92369	0.24667		
C. Total	478	148.23704			

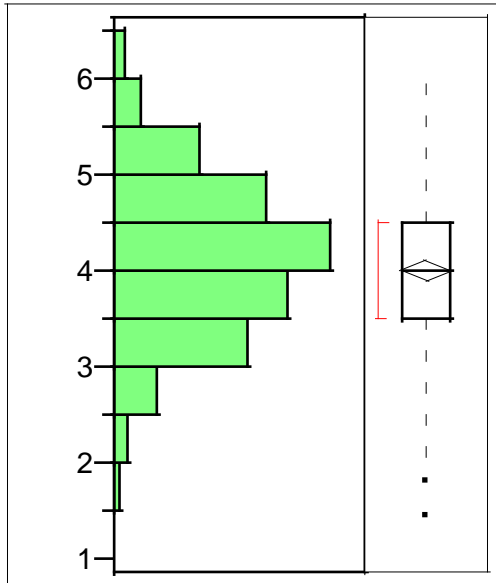
Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
B	95	1.87368	0.05096	1.7736	1.9738
G	96	2.65625	0.05069	2.5566	2.7559
O	96	2.25521	0.05069	2.1556	2.3548
R	96	2.09375	0.05069	1.9941	2.1934
Y	96	2.22292	0.05069	2.1233	2.3225

Std Error uses a pooled estimate of error variance

Table 8 Interrater reliabilities (To be provided)

Table 9
Distributions
MeanOfA1



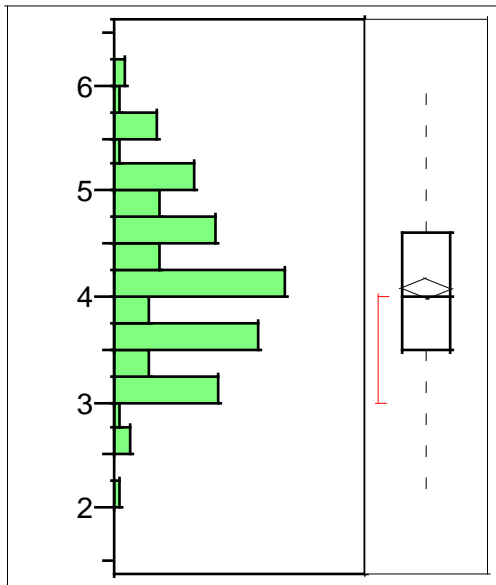
Quantiles

100.0%	maximum	6.0000
99.5%		6.0000
97.5%		5.6463
90.0%		5.0000
75.0%	quartile	4.5000
50.0%	median	4.0000
25.0%	quartile	3.5000
10.0%		3.0000
2.5%		2.0125
0.5%		1.5718
0.0%	minimum	1.5000

Moments

Mean	3.9920833
Std Dev	0.837697
Std Err Mean	0.0540731
upper 95% Mean	4.0986041
lower 95% Mean	3.8855626
N	240

Table 10
MeanOfA2



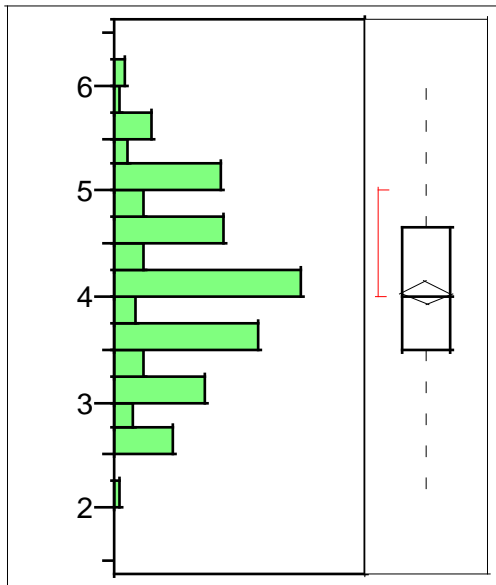
Quantiles

100.0%	maximum	6.0000
99.5%		6.0000
97.5%		5.5975
90.0%		5.0000
75.0%	quartile	4.6125
50.0%	median	4.0000
25.0%	quartile	3.5000
10.0%		3.0000
2.5%		2.5037
0.5%		2.0000
0.0%	minimum	2.0000

Moments

Mean	4.0752083
Std Dev	0.7941074
Std Err Mean	0.0512594
upper 95% Mean	4.1761863
lower 95% Mean	3.9742304
N	240

Table 11
MeanOfB1



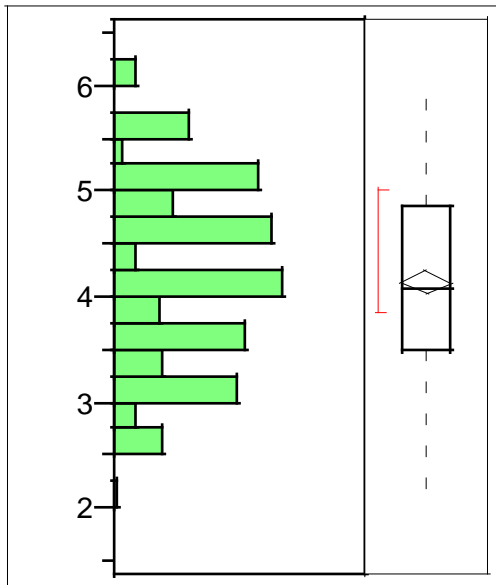
Quantiles

100.0%	maximum	6.0000
99.5%		6.0000
97.5%		5.5000
90.0%		5.0000
75.0%	quartile	4.6500
50.0%	median	4.0000
25.0%	quartile	3.5000
10.0%		3.0000
2.5%		2.5000
0.5%		2.0000
0.0%	minimum	2.0000

Moments

Mean	4.0307708
Std Dev	0.8594429
Std Err Mean	0.0554768
upper 95% Mean	4.1400568
lower 95% Mean	3.9214849
N	240

Table 12
MeanOfB2



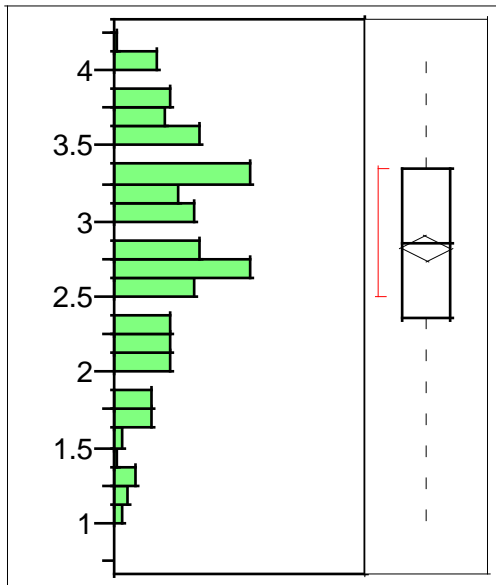
Quantiles

100.0%	maximum	6.0000
99.5%		6.0000
97.5%		5.6500
90.0%		5.3300
75.0%	quartile	4.8500
50.0%	median	4.0750
25.0%	quartile	3.5000
10.0%		3.0000
2.5%		2.5000
0.5%		2.1025
0.0%	minimum	2.0000

Moments

Mean	4.1383333
Std Dev	0.8739305
Std Err Mean	0.056412
upper 95% Mean	4.2494615
lower 95% Mean	4.0272052
N	240

Table 13
MeanOfGlobal



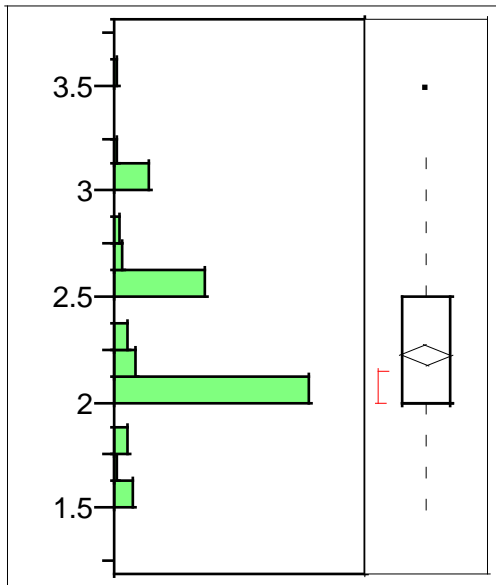
Quantiles

100.0%	maximum	4.1500
99.5%		4.1193
97.5%		4.0000
90.0%		3.7000
75.0%	quartile	3.3500
50.0%	median	2.8500
25.0%	quartile	2.3500
10.0%		1.8050
2.5%		1.3012
0.5%		1.0000
0.0%	minimum	1.0000

Moments

Mean	2.8122917
Std Dev	0.7138242
Std Err Mean	0.0460772
upper 95% Mean	2.9030609
lower 95% Mean	2.7215225
N	240

Table 14
MeanOfImprovement



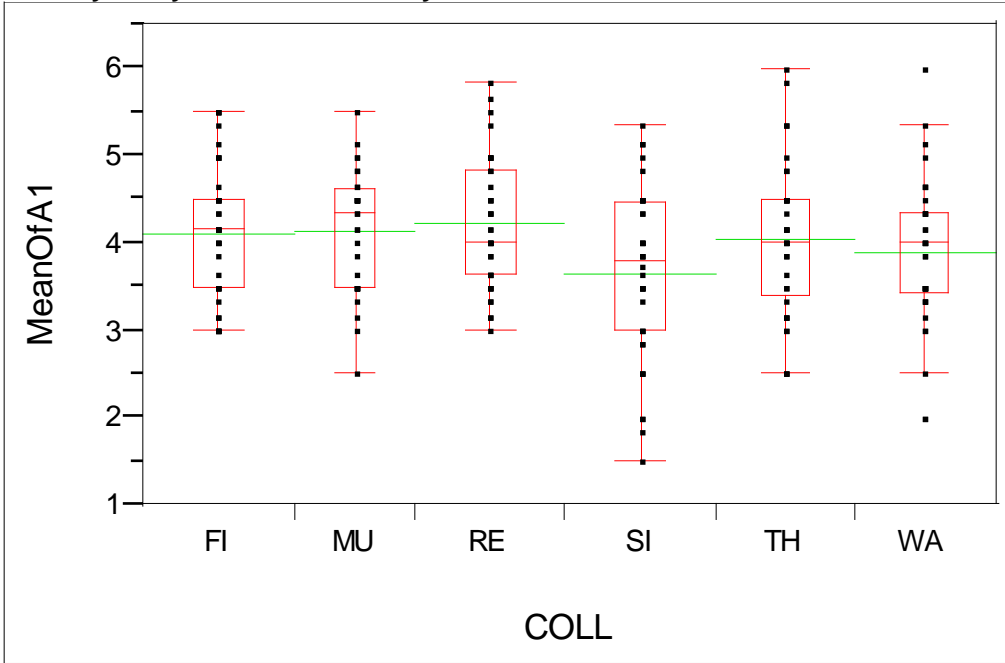
Quantiles

100.0%	maximum	3.5000
99.5%		3.4282
97.5%		3.0000
90.0%		2.8500
75.0%	quartile	2.5000
50.0%	median	2.0000
25.0%	quartile	2.0000
10.0%		2.0000
2.5%		1.5000
0.5%		1.5000
0.0%	minimum	1.5000

Moments

Mean	2.220625
Std Dev	0.3822987
Std Err Mean	0.0246773
upper 95% Mean	2.2692377
lower 95% Mean	2.1720123
N	240

Table 15
Oneway Analysis of MeanOfA1 By COLL



Oneway Anova

Analysis of Variance

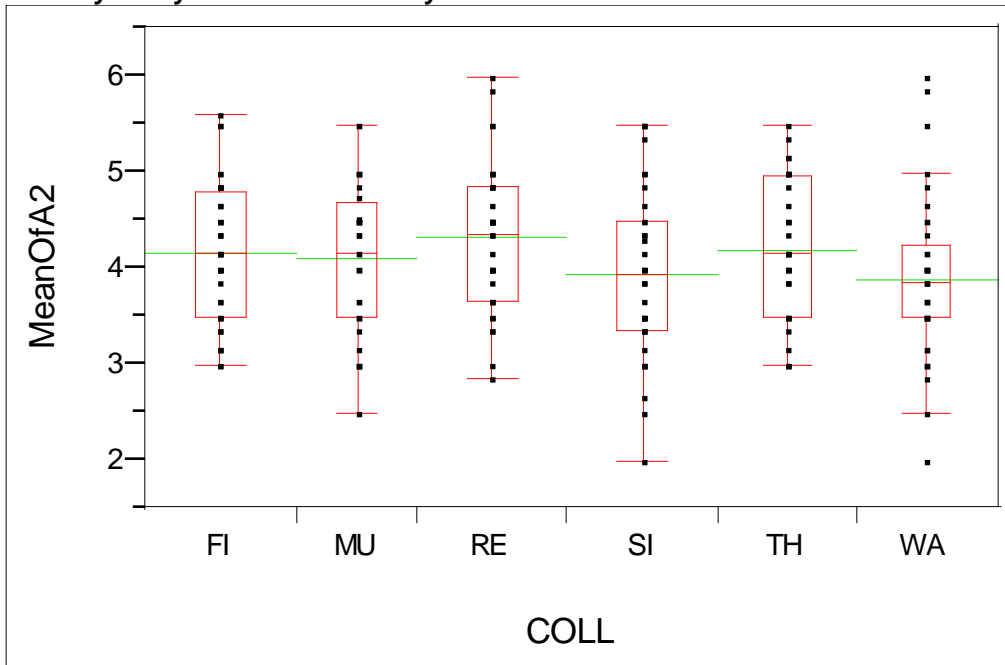
Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
COLL	5	8.20135	1.64027	2.4207	0.0368
Error	218	147.72024	0.67762		
C. Total	223	155.92159			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
FI	40	4.09375	0.13016	3.8372	4.3503
MU	32	4.14531	0.14552	3.8585	4.4321
RE	39	4.21026	0.13181	3.9505	4.4700
SI	40	3.65125	0.13016	3.3947	3.9078
TH	36	4.03750	0.13720	3.7671	4.3079
WA	37	3.87973	0.13533	3.6130	4.1465

Std Error uses a pooled estimate of error variance

Table 16
Oneway Analysis of MeanOfA2 By COLL



Oneway Anova

Analysis of Variance

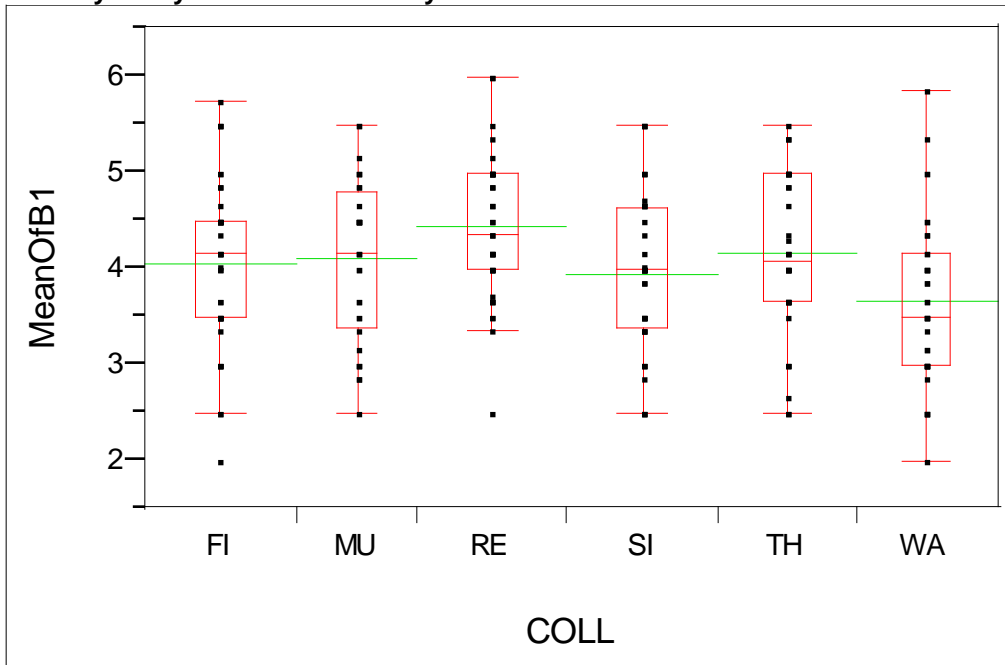
Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
COLL	5	5.10086	1.02017	1.6676	0.1436
Error	218	133.36350	0.61176		
C. Total	223	138.46436			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
FI	40	4.15625	0.12367	3.9125	4.4000
MU	32	4.09063	0.13827	3.8181	4.3631
RE	39	4.31410	0.12524	4.0673	4.5609
SI	40	3.93000	0.12367	3.6863	4.1737
TH	36	4.18889	0.13036	3.9320	4.4458
WA	37	3.88243	0.12858	3.6290	4.1359

Std Error uses a pooled estimate of error variance

Table 17
Oneway Analysis of MeanOfB1 By COLL



Oneway Anova

Analysis of Variance

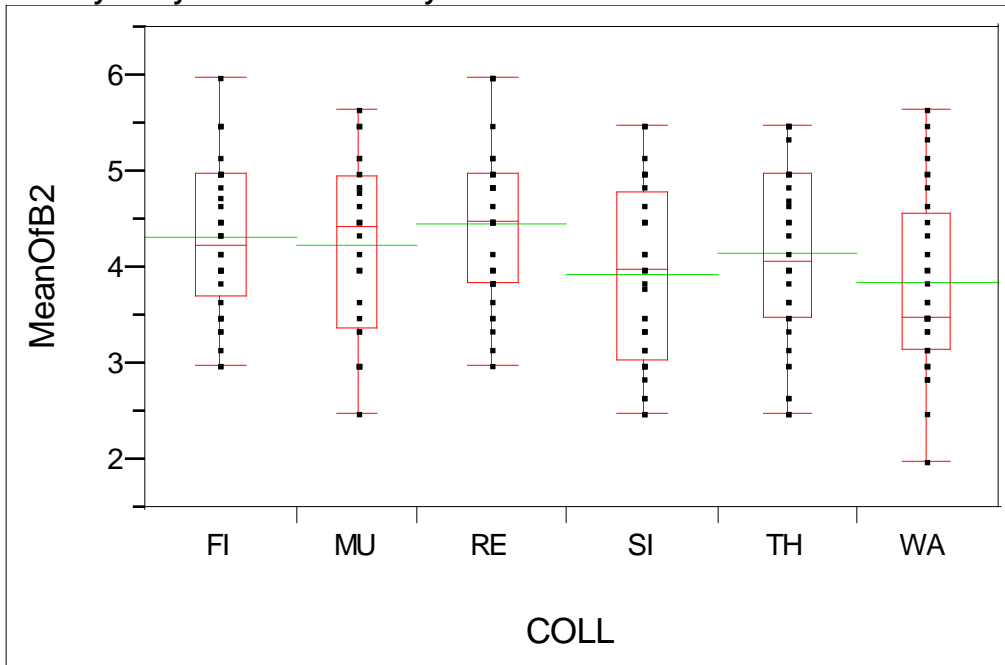
Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
COLL	5	12.77667	2.55533	3.7267	0.0029
Error	218	149.47829	0.68568		
C. Total	223	162.25495			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
FI	40	4.04250	0.13093	3.7845	4.3005
MU	32	4.09219	0.14638	3.8037	4.3807
RE	39	4.43462	0.13260	4.1733	4.6959
SI	40	3.94375	0.13093	3.6857	4.2018
TH	36	4.14028	0.13801	3.8683	4.4123
WA	37	3.64054	0.13613	3.3722	3.9088

Std Error uses a pooled estimate of error variance

Table 18
Oneway Analysis of MeanOfB2 By COLL



Oneway Anova

Analysis of Variance

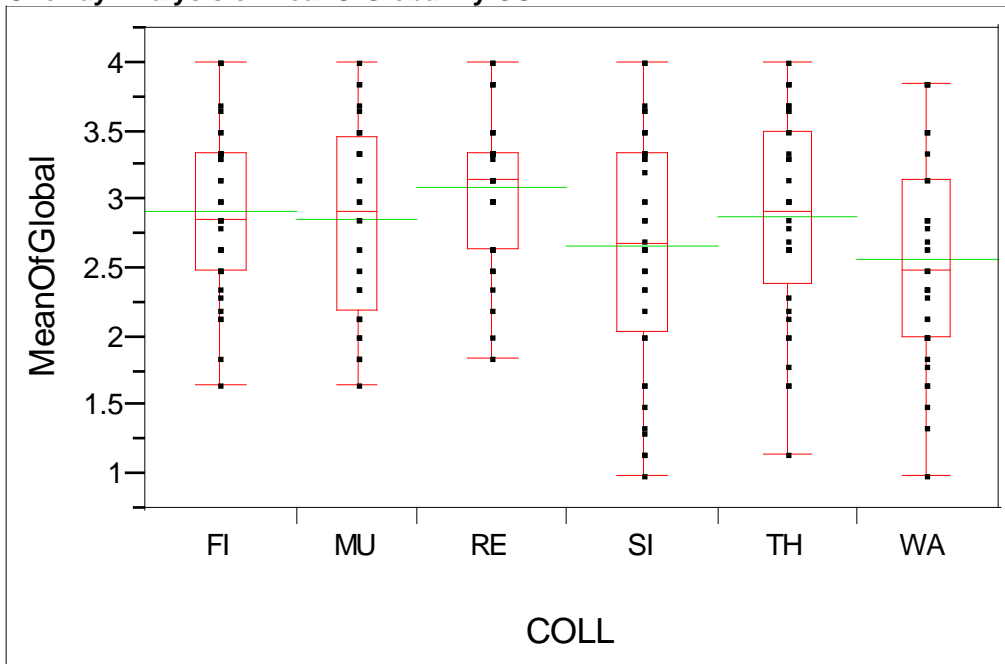
Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
COLL	5	10.37520	2.07504	2.8480	0.0163
Error	218	158.83532	0.72860		
C. Total	223	169.21053			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
FI	40	4.30625	0.13496	4.0403	4.5722
MU	32	4.23281	0.15089	3.9354	4.5302
RE	39	4.46282	0.13668	4.1934	4.7322
SI	40	3.94000	0.13496	3.6740	4.2060
TH	36	4.15278	0.14226	3.8724	4.4332
WA	37	3.83784	0.14033	3.5613	4.1144

Std Error uses a pooled estimate of error variance

Table 19
Oneway Analysis of MeanOfGlobal By COLL



Oneway Anova

Analysis of Variance

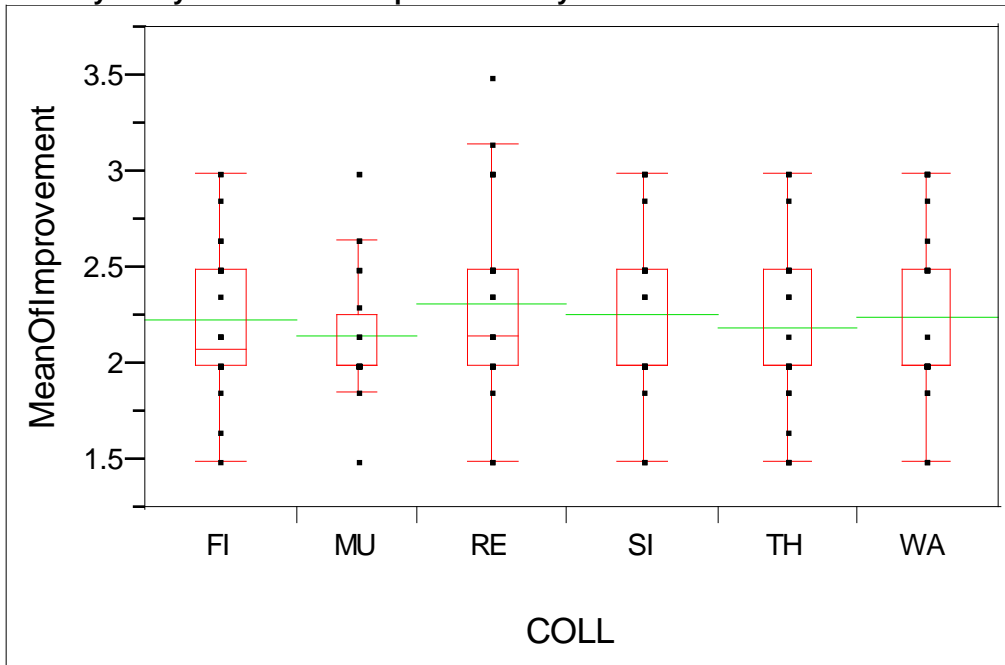
Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
COLL	5	6.87753	1.37551	2.8836	0.0153
Error	218	103.98657	0.47700		
C. Total	223	110.86410			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
FI	40	2.92750	0.10920	2.7123	3.1427
MU	32	2.86719	0.12209	2.6266	3.1078
RE	39	3.08718	0.11059	2.8692	3.3051
SI	40	2.65875	0.10920	2.4435	2.8740
TH	36	2.87500	0.11511	2.6481	3.1019
WA	37	2.56351	0.11354	2.3397	2.7873

Std Error uses a pooled estimate of error variance

Table 20
Oneway Analysis of MeanOfImprovement By COLL



Oneway Anova

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
COLL	5	0.618894	0.123779	0.8164	0.5391
Error	218	33.050381	0.151607		
C. Total	223	33.669275			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
FI	40	2.22750	0.06156	2.1062	2.3488
MU	32	2.14844	0.06883	2.0128	2.2841
RE	39	2.31538	0.06235	2.1925	2.4383
SI	40	2.26000	0.06156	2.1387	2.3813
TH	36	2.18194	0.06489	2.0540	2.3098
WA	37	2.23919	0.06401	2.1130	2.3654

Std Error uses a pooled estimate of error variance

Table 21
Multivariate
Correlations

	MeanOfGlobal	MeanOfImprove ment	Course 1 Grade	Course 2 Grade
MeanOfGlobal	1.0000	0.2923	0.3523	0.3117
MeanOfImprove ment	0.2923	1.0000	0.0738	0.0463
Course 1 Grade	0.3523	0.0738	1.0000	0.5437
Course 2 Grade	0.3117	0.0463	0.5437	1.0000

2 rows not used due to missing values.

Table 22
Multivariate
Correlations

	MeanOfGlobal
MeanOfGlobal	1.0000
FATHER ED	0.2621
MOTHER ED	0.2852
INCOME	0.2528
HS_GPA	-0.0175
SATVOFL	0.2630
SATMOFL	0.0958

Table 23
Multivariate
Correlations

	MeanOfImprove ment
MeanOfImprove ment	1.0000
FATHER ED	0.1656
MOTHER ED	0.1713
INCOME	0.0334
HS_GPA	-0.0584
SATVOFL	-0.0080
SATMOFL	0.0819

Table 24
Multivariate
Correlations

	Course 1 Grade	Course 2 Grade
Course 1 Grade	1.0000	0.5727
Course 2 Grade	0.5727	1.0000
FATHER ED	0.2008	0.2577
MOTHER ED	0.1700	0.2034
INCOME	0.2224	0.2051
HS_GPA	0.1135	0.1034
SATVOFL	0.3209	0.3027
SATMOFL	0.0687	0.1270

Appendix A: Guidelines to reviewers

Guidelines for Readers of “University Writing” samples.

[These guidelines are adapted from the CEP resolution dated May 27, 2003.]

Each portfolio will consist of four papers which correspond to the specific writing assignments of the individual writing program. These four papers were to be an early and late paper assigned to fulfill the requirements of the two quarter mandatory university writing program required system-wide. The papers should be in chronological order. The individual portfolios will contain the four writing samples as well as the prompts that were assigned. The prompts for a given writing program may be very similar across all students or may vary considerably according to which section a student was enrolled in. The communality of the writing prompts differs by program.

You have a master grading sheet that should be used to record your grades. The sheet has a row for each of the portfolios you are to read. There are, for each portfolio, six columns that you will need to fill out. The first four are the “subject A scores” for each of the papers. The fifth score is a global A-F “grade” that you would assign to the portfolio as a whole based upon your assessment of both the sentence level and rhetorical level of writing. The sixth score is four point score based upon your judgment of the degree of improvement you sense over the four writing samples.

Subject A Scores – A six point score assigned to each paper

A **6** paper commands attention because of its insightful development and mature style. It presents a cogent response to the text, elaborating that response with well-chosen examples and persuasive reasoning. The 6 paper shows that its writer can usually choose words aptly, use sophisticated sentences effectively, and observe the conventions of written English.

A **5** paper is clearly competent. It presents a thoughtful response to the text, elaborating that response with appropriate examples and sensible reasoning. A 5 paper typically has a less fluent and complex style than a 6, but does show that its writer can usually choose words accurately, vary sentences effectively, and observe the conventions of written English.

A **4** paper is satisfactory, sometimes marginally so. It presents an adequate response to the text, elaborating that response with sufficient examples and acceptable reasoning. Just as these examples and this reasoning will ordinarily be less developed than those in 5 papers, so will the 4 paper's style be less effective. Nevertheless, a 4 paper shows that its writer can usually choose words of sufficient precision, control sentences of reasonable variety, and observe the conventions of written English.

A **3** paper is unsatisfactory in one or more of the following ways. It may respond to the text illogically; it may lack coherent structure or elaboration with examples; it may reflect an incomplete understanding of the text or the topic. Its prose is usually characterized by at least one of the following: frequently imprecise word choice; little sentence variety; occasional major errors in grammar and usage, or frequent minor errors.

A **2** paper shows serious weaknesses, ordinarily of several kinds. It frequently presents a simplistic, inappropriate, or incoherent response to the text, one that may suggest some significant

misunderstanding of the text or the topic. Its prose is usually characterized by at least one of the following: simplistic or inaccurate word choice; monotonous or fragmented sentence structure; many repeated errors in grammar and usage.

A 1 paper suggests severe difficulties in reading and writing conventional English. It may disregard the topic's demands, or it may lack any appropriate pattern of structure or development. It may be inappropriately brief. It often has a pervasive pattern of errors in word choice, sentence structure, grammar, and usage.

A Global “A-F” grade - a single score assigned to the entire portfolio

This single letter grade is assigned to the portfolio as a whole. It is a global judgment of the quality of the writing. The grader should take the following perspective into consideration when assigning the grade.

In looking at student papers, committee members should concern themselves both with the sentence level and with the rhetorical level of student essays.

On the sentence level, they should look for serious problems of usage and grammar, such as

- lack of subject-verb and pronoun-referent agreement
- dangling modifiers
- comma splices
- sentence fragments
- run-on sentences, run-on paragraphs
- redundant or inappropriate use of words
- stilted language
- overly informal language
- jargon
- vagueness

They should also look for examples of good writing on the sentence level, such as:

- clear and precise word choice
- appropriate tone
- good use of subordination, apposition, and other structural devices
- effective placement of emphasis; verbal emphasis on what is intellectually important
- effective embedding of quotations
- smooth transitions within and among sentences

On the rhetorical level, committee members should look for such serious problems as:

- absence of thesis or direction
- unclear relationship between the various parts of the paper and the thesis
- lack of rhetorical emphasis (i.e., clear verbal discrimination of central and peripheral issues)
- poor use of evidence or logic
- over-lengthy introductions and conclusions, irrelevant argument or quotation, excess verbiage in general

They should also look for examples of good writing on the rhetorical level, such as:

precise statement of thesis
effective choice and arrangement of evidence
precision of organization, with easily recognizable connections among sections and paragraphs
steady, accessible development of an argument; avoidance of anticlimax
elimination of all that is not useful to the main argument
ability to distinguish substantial from insubstantial intellectual issues
recognition of plausible counterarguments, and effective response to them
effective use of irony, humor, and other means of eliciting emotion

Papers should never be favorably assessed simply because they are “original” or because they take the “correct” position on the subjects they address. An “original” paper is sometimes a very poor one, and the assessment must always concern writing quality, not opinions per se.

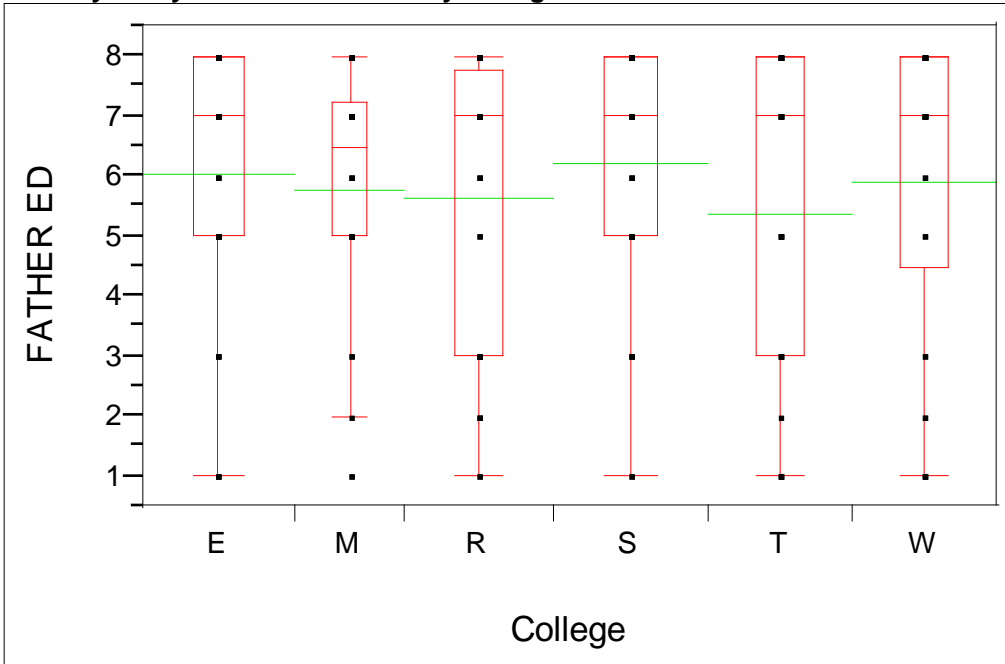
Degree of Improvement Score - a single score based upon your judgment of the degree to which writing seems to have improved (or worsened) over the two quarter span.

- 4 - a great improvement in the writing
- 3 - some improvement in the writing
- 2 - no improvement, writing is essentially unchanged
- 1 - a noticeable lessening in the quality of the writing

This score is a very subjective judgment on your part.

Appendix B: Program Differences

Oneway Analysis of FATHER ED By College



Oneway Anova

Analysis of Variance

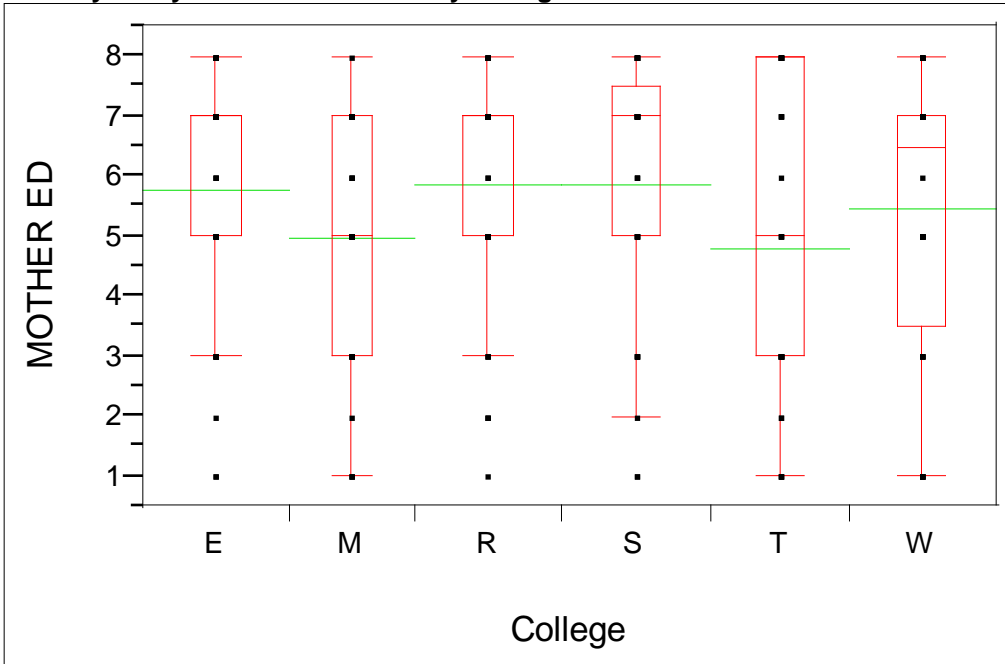
Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
College	5	15.2017	3.04035	0.5663	0.7258
Error	199	1068.4763	5.36923		
C. Total	204	1083.6780			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
E	37	6.02703	0.38094	5.2758	6.7782
M	26	5.76923	0.45443	4.8731	6.6654
R	36	5.63889	0.38619	4.8773	6.4004
S	37	6.18919	0.38094	5.4380	6.9404
T	35	5.37143	0.39167	4.5991	6.1438
W	34	5.91176	0.39739	5.1281	6.6954

Std Error uses a pooled estimate of error variance

Oneway Analysis of MOTHER ED By College



Oneway Anova

Analysis of Variance

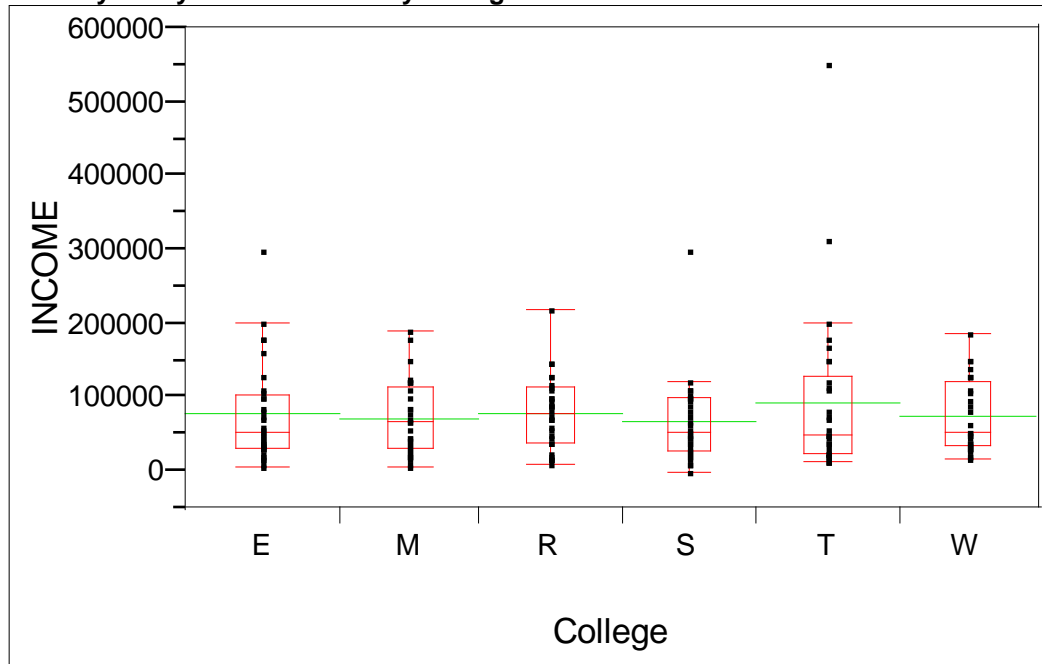
Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
College	5	38.5584	7.71168	1.5706	0.1698
Error	207	1016.3524	4.90991		
C. Total	212	1054.9108			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
E	37	5.78378	0.36428	5.0656	6.5020
M	31	4.96774	0.39798	4.1831	5.7523
R	37	5.83784	0.36428	5.1197	6.5560
S	37	5.83784	0.36428	5.1197	6.5560
T	35	4.77143	0.37454	4.0330	5.5098
W	36	5.44444	0.36931	4.7164	6.1725

Std Error uses a pooled estimate of error variance

Oneway Analysis of INCOME By College



Oneway Anova

Analysis of Variance

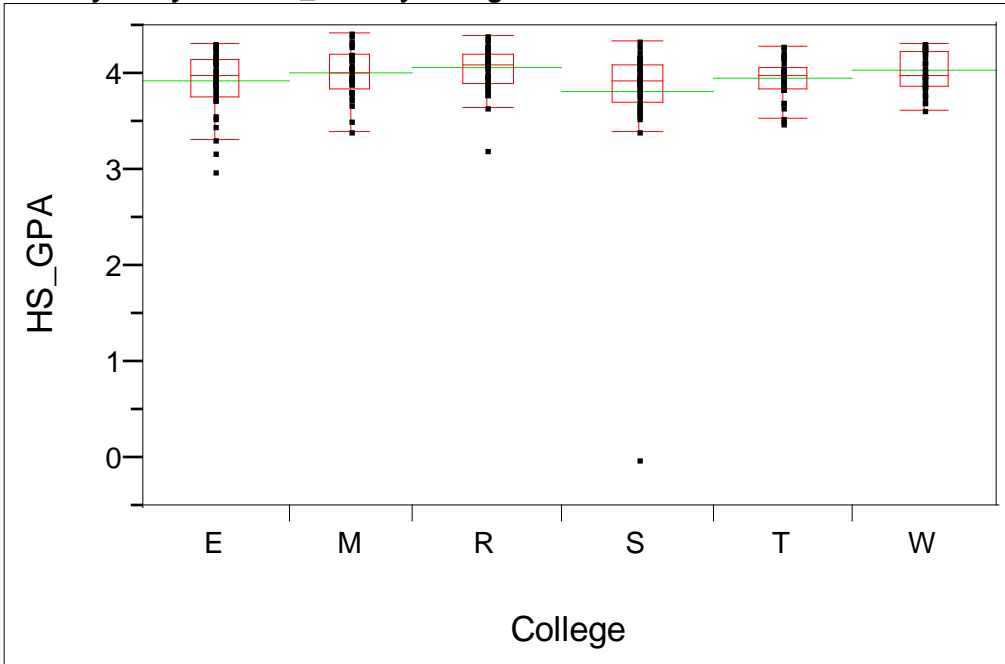
Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
College	5	1.22014e10	2.44027e9	0.5315	0.7522
Error	173	7.94296e11	4.59131e9		
C. Total	178	8.06497e11			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
E	33	76410.5	11795	53129	99692
M	29	72048.7	12583	47214	96884
R	30	79256.7	12371	54839	103674
S	28	67521.1	12805	42246	92796
T	30	93890.1	12371	69472	118308
W	29	73945.6	12583	49111	98781

Std Error uses a pooled estimate of error variance

Oneway Analysis of HS_GPA By College



Oneway Anova

Analysis of Variance

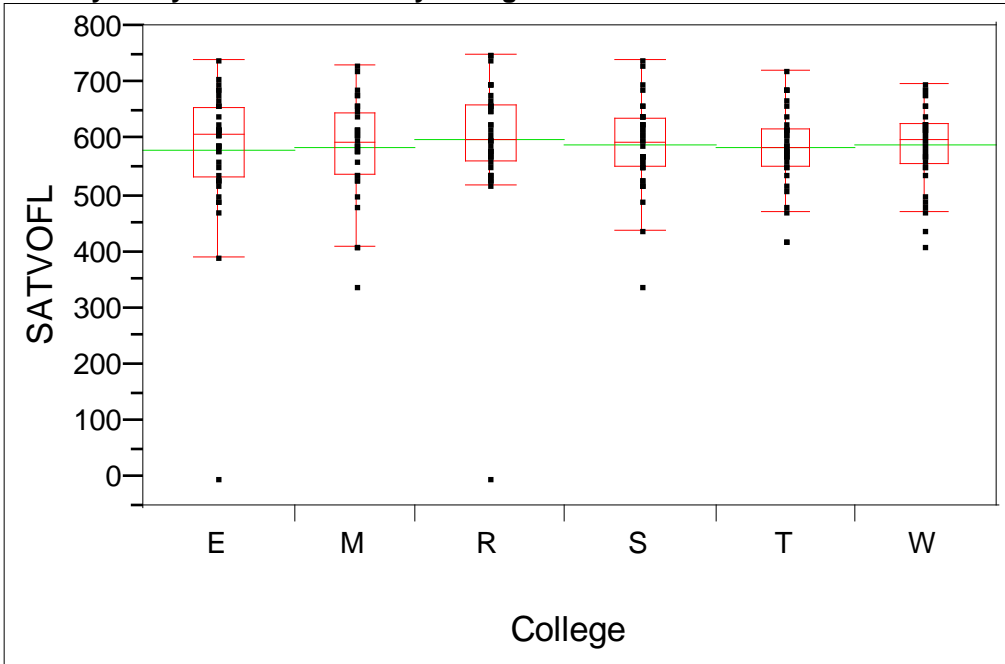
Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
College	5	1.407531	0.281506	2.2209	0.0534
Error	215	27.252216	0.126754		
C. Total	220	28.659747			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
E	38	3.92868	0.05776	3.8148	4.0425
M	32	4.01469	0.06294	3.8906	4.1387
R	39	4.05795	0.05701	3.9456	4.1703
S	39	3.82410	0.05701	3.7117	3.9365
T	36	3.96250	0.05934	3.8455	4.0795
W	37	4.03270	0.05853	3.9173	4.1481

Std Error uses a pooled estimate of error variance

Oneway Analysis of SATVOFL By College



Oneway Anova

Analysis of Variance

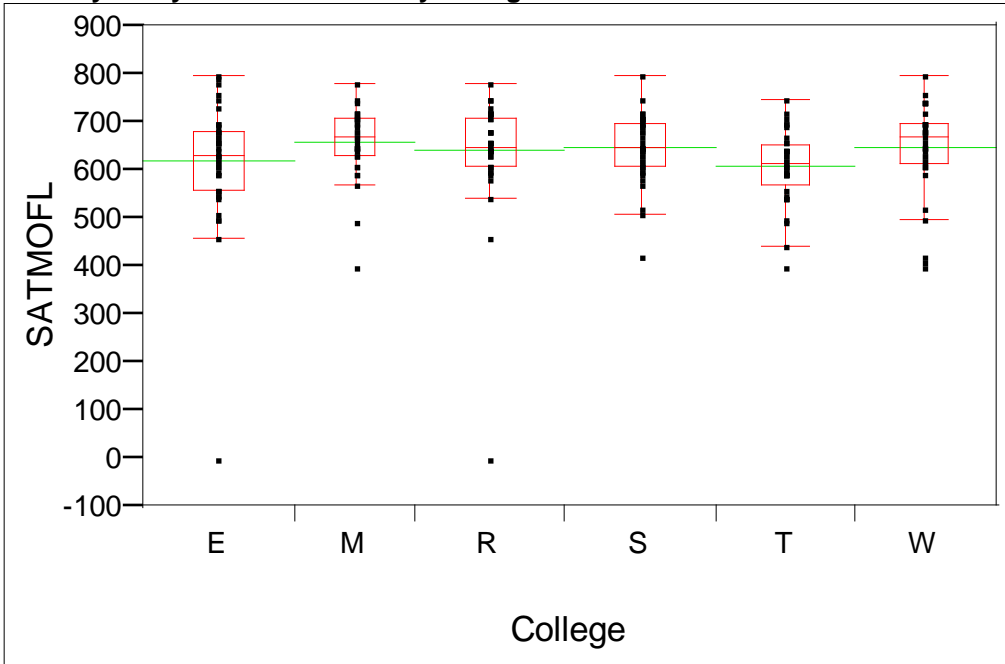
Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
College	5	6847.7	1369.55	0.1573	0.9776
Error	218	1897573.7	8704.47		
C. Total	223	1904421.4			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
E	40	582.500	14.752	553.43	611.57
M	32	585.938	16.493	553.43	618.44
R	39	599.487	14.940	570.04	628.93
S	40	589.750	14.752	560.68	618.82
T	36	585.556	15.550	554.91	616.20
W	37	590.811	15.338	560.58	621.04

Std Error uses a pooled estimate of error variance

Oneway Analysis of SATMOFL By College



Oneway Anova

Analysis of Variance

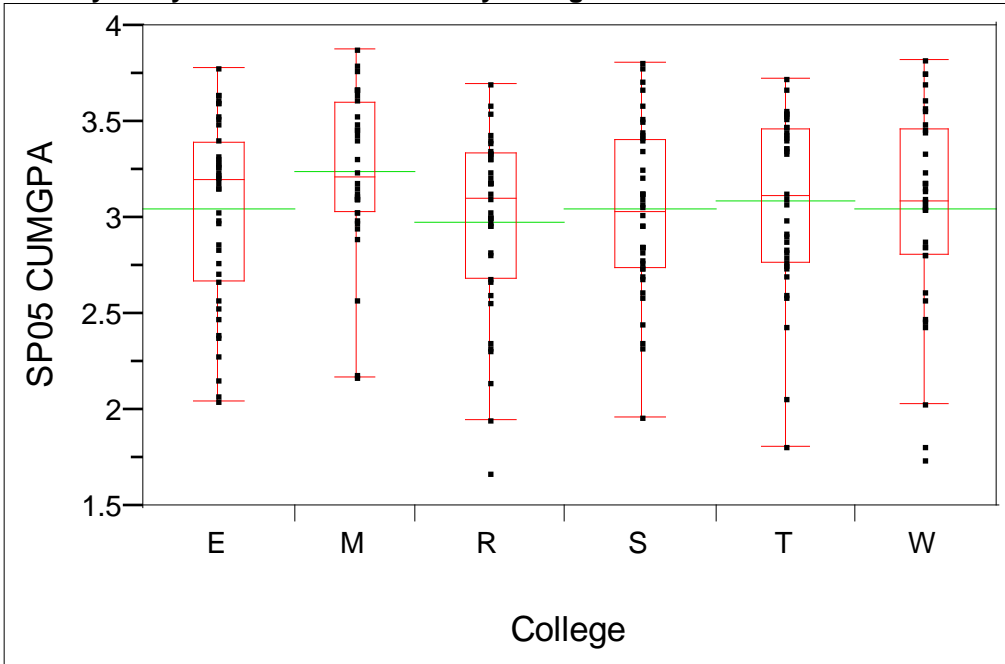
Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
College	5	67251.9	13450.4	1.3900	0.2291
Error	218	2109560.6	9676.9		
C. Total	223	2176812.5			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
E	40	619.250	15.554	588.59	649.91
M	32	658.750	17.390	624.48	693.02
R	39	640.256	15.752	609.21	671.30
S	40	648.750	15.554	618.09	679.41
T	36	608.611	16.395	576.30	640.92
W	37	648.108	16.172	616.23	679.98

Std Error uses a pooled estimate of error variance

Oneway Analysis of SP05 CUMGPA By College



Oneway Anova

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
College	5	1.359927	0.271985	1.2753	0.2756
Error	218	46.491609	0.213264		
C. Total	223	47.851535			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
E	40	3.04728	0.07302	2.9034	3.1912
M	32	3.24554	0.08164	3.0846	3.4064
R	39	2.98514	0.07395	2.8394	3.1309
S	40	3.04298	0.07302	2.8991	3.1869
T	36	3.08955	0.07697	2.9379	3.2412
W	37	3.04180	0.07592	2.8922	3.1914

Std Error uses a pooled estimate of error variance

Contingency Analysis of USUBJA By College

Contingency Table

College By USUBJA

Count Total % Col % Row %	A	B	C	N	U	X	
E	13 5.80 24.53 32.50	9 4.02 15.52 22.50	0 0.00 0.00 0.00	1 0.45 12.50 2.50	12 5.36 20.34 30.00	5 2.23 13.89 12.50	40 17.86
M	5 2.23 9.43 15.63	12 5.36 20.69 37.50	4 1.79 40.00 12.50	1 0.45 12.50 3.13	5 2.23 8.47 15.63	5 2.23 13.89 15.63	32 14.29
R	11 4.91 20.75 28.21	13 5.80 22.41 33.33	1 0.45 10.00 2.56	2 0.89 25.00 5.13	8 3.57 13.56 20.51	4 1.79 11.11 10.26	39 17.41
S	9 4.02 16.98 22.50	9 4.02 15.52 22.50	4 1.79 40.00 10.00	0 0.00 0.00 0.00	11 4.91 18.64 27.50	7 3.13 19.44 17.50	40 17.86
T	7 3.13 13.21 19.44	7 3.13 12.07 19.44	1 0.45 10.00 2.78	1 0.45 12.50 2.78	15 6.70 25.42 41.67	5 2.23 13.89 13.89	36 16.07
W	8 3.57 15.09 21.62	8 3.57 13.79 21.62	0 0.00 0.00 0.00	3 1.34 37.50 8.11	8 3.57 13.56 21.62	10 4.46 27.78 27.03	37 16.52
	53 23.66	58 25.89	10 4.46	8 3.57	59 26.34	36 16.07	224

Tests

Source	DF	-LogLike	RSquare (U)
Model	25	16.54298	0.0463
Error	194	340.49179	
C. Total	219	357.03478	
N	224		

Test	ChiSquare	Prob>ChiSq
Likelihood Ratio	33.086	0.1289
Pearson	31.686	0.1673

Warning: 20% of cells have expected count less than 5, ChiSquare suspect

Contingency Analysis of GNDR By College
Contingency Table
 College By GNDR

Count	F	M	
Total %			
Col %			
Row %			
E	29	11	40
	12.95	4.91	17.86
	22.48	11.58	
	72.50	27.50	
M	17	15	32
	7.59	6.70	14.29
	13.18	15.79	
	53.13	46.88	
R	24	15	39
	10.71	6.70	17.41
	18.60	15.79	
	61.54	38.46	
S	23	17	40
	10.27	7.59	17.86
	17.83	17.89	
	57.50	42.50	
T	24	12	36
	10.71	5.36	16.07
	18.60	12.63	
	66.67	33.33	
W	12	25	37
	5.36	11.16	16.52
	9.30	26.32	
	32.43	67.57	
	129	95	224
	57.59	42.41	

Tests

Source	DF	-LogLike	RSquare (U)
Model	5	7.54295	0.0494
Error	218	145.13166	
C. Total	223	152.67461	
N	224		

Test	ChiSquare	Prob>ChiSq
Likelihood Ratio	15.086	0.0100
Pearson	14.953	0.0106

Guidelines for Readers of “University Writing” samples.

[These guidelines are adapted from the CEP resolution dated May 27, 2003.]

Each portfolio will consist of four papers which correspond to the specific writing assignments of the individual writing program. These four papers were to be an early and late paper assigned in courses that fulfill the requirements of the two-quarter mandatory university writing program. The papers should be in chronological order and labeled A1, A2, B1, and B2. The individual portfolios will contain the four writing samples as well as the prompts that were assigned. The prompts for a given writing program may be very similar across all students or may vary considerably according to which section a student was enrolled in. The communality of the writing prompts differs by program.

You have a master grading sheet that should be used to record your grades. The sheet has a row for each of the portfolios you are to read. There are, for each portfolio, six columns that you will need to fill out. The first four are the “subject A scores” for each of the papers. The fifth score is a global A-F “grade” that you would assign to the portfolio as a whole based upon your assessment of both the sentence level and rhetorical level of writing. The sixth score is four point score based upon your judgment of the degree of improvement you sense over the four writing samples.

Subject A Scores – A six-point score assigned to each paper

A **6** paper commands attention because of its insightful development and mature style. It presents a cogent response to the text, elaborating that response with well-chosen examples and persuasive reasoning. The 6 paper shows that its writer can usually choose words aptly, use sophisticated sentences effectively, and observe the conventions of written English.

A **5** paper is clearly competent. It presents a thoughtful response to the text, elaborating that response with appropriate examples and sensible reasoning. A 5 paper typically has a less fluent and complex style than a 6, but does show that its writer can usually choose words accurately, vary sentences effectively, and observe the conventions of written English.

A **4** paper is satisfactory, sometimes marginally so. It presents an adequate response to the text, elaborating that response with sufficient examples and acceptable reasoning. Just as these examples and this reasoning will ordinarily be less developed than those in 5 papers, so will the 4 paper's style be less effective. Nevertheless, a 4 paper shows that its writer can usually choose words of sufficient precision, control sentences of reasonable variety, and observe the conventions of written English.

A **3** paper is unsatisfactory in one or more of the following ways. It may respond to the text illogically; it may lack coherent structure or elaboration with examples; it may reflect an incomplete understanding of the text or the topic. Its prose is usually characterized by at least one of the following: frequently imprecise word choice; little sentence variety; occasional major errors in grammar and usage, or frequent minor errors.

A **2** paper shows serious weaknesses, ordinarily of several kinds. It frequently presents a simplistic, inappropriate, or incoherent response to the text, one that may suggest some significant misunderstanding of the text or the topic. Its prose is usually characterized by at least one of the following: simplistic or inaccurate word choice; monotonous or fragmented sentence structure; many repeated errors in grammar and usage.

7/21/2009

1:44:19 PM

P:\UG Educ\Writing Assessments\Assessment of Freshman Writing_04-05
data\Guidelines for Readers of Writing SamplesRev.doc

A 1 paper suggests severe difficulties in reading and writing conventional English. It may disregard the topic's demands, or it may lack any appropriate pattern of structure or development. It may be inappropriately brief. It often has a pervasive pattern of errors in word choice, sentence structure, grammar, and usage.

A Global “A-F” grade - a single score assigned to the entire portfolio

This single letter grade is assigned to the portfolio as a whole. It is a global judgment of the quality of the writing. The grader should take the following perspective into consideration when assigning the grade.

In looking at student papers, committee members should concern themselves both with the sentence level and with the rhetorical level of student essays.

On the sentence level, graders should look for serious problems of usage and grammar, such as:

- lack of subject-verb and pronoun-referent agreement
- dangling modifiers
- comma splices
- sentence fragments
- run-on sentences, run-on paragraphs
- redundant or inappropriate use of words
- stilted language
- overly informal language
- jargon
- vagueness

They should also look for examples of good writing on the sentence level, such as:

- clear and precise word choice
- appropriate tone
- good use of subordination, apposition, and other structural devices
- effective placement of emphasis; verbal emphasis on what is intellectually important
- effective embedding of quotations
- smooth transitions within and among sentences

On the rhetorical level, committee members should look for such serious problems as:

- absence of thesis or direction
- unclear relationship between the various parts of the paper and the thesis
- lack of rhetorical emphasis (i.e., clear verbal discrimination of central and peripheral issues)
- poor use of evidence or logic
- over-lengthy introductions and conclusions, irrelevant argument or quotation, excess verbiage in general

They should also look for examples of good writing on the rhetorical level, such as:

- precise statement of thesis

effective choice and arrangement of evidence
precision of organization, with easily recognizable connections among sections and paragraphs
steady, accessible development of an argument; avoidance of anticlimax
elimination of all that is not useful to the main argument
ability to distinguish substantial from insubstantial intellectual issues
recognition of plausible counterarguments, and effective response to them
effective use of irony, humor, and other means of eliciting emotion

Papers should never be favorably assessed simply because they are “original” or because they take the “correct” position on the subjects they address. An “original” paper is sometimes a very poor one, and the assessment must always concern writing quality, not opinions per se.

Degree of Improvement Score - a single score based upon your judgment of the degree to which writing seems to have improved (or worsened) over the two-quarter span:

- 4 - a great improvement in the writing
- 3 - some improvement in the writing
- 2 - no improvement, writing is essentially unchanged
- 1 - a noticeable lessening in the quality of the writing

This score is a subjective judgment on your part.

Freshman Writing Evaluation
May 2006
Minority Report

Received
MAY 08 2006
Academic Senate
UCSD

Paul A. Libby

I have no trouble with the description in the majority report of the procedure used by the five faculty members in this exercise nor with the finding that the average rating over all readers, over all colleges and over all samples corresponds to a "satisfactory" quality of writing by our freshman nor to the discussion of "...the variations in outcome among college." I point out later for several minor points of concern in the majority report but my main concern is its reticent tone.. The majority seem to take the view that the faculty readers are to report their findings and to leave to others, presumably CEP and the campus administration, their interpretation and significance. I do not accept this narrow view of our responsibilities.

Some but not all of the readers were briefed at the outset of this exercise on the previous history of studies of the freshman writing programs at UCSD. That history extends back to the 1970's and consists of periodic exercises examining the effectiveness of these programs. Indeed I participated in a review of the Warren College program five or so years ago at the same time that other faculty teams reviewed the programs of the other colleges. I know that the report of our committee was highly critical of the Warren program and I understand that the reports of the faculty groups assigned the other colleges were equally critical. At the time CEP took no action on these findings. In the briefing in question we were told that although these and earlier reviews of freshman writing had been highly critical, CEP believed it needed "data" before it could take any remedial steps. The present exercise was intended to provide these data.

As a result of our reading we now know several important features about the writing abilities of our freshman and about the effectiveness of the various writing programs in improving those abilities. Only half can write in a satisfactory manner. We also know that students who write well at the beginning of the first quarter of the program write well at the end of the second quarter and the poor writers at the beginning of the first quarter are poor writers at the end of the second quarter. In brief the writing programs do not improve the writing skills of our freshman.

My problem with the majority report is that it fails to make these statements unequivocally. Indeed under the section "Concerning the 'improvement finding'" it discusses possible reasons for this finding to be spurious. I consider these findings to be incontrovertible.

I now address my minor concerns with the majority report. Emphasized there is that averaging over all readers and all students and all colleges results in a global grade of between a B- and B and hence to a statement that the overall writing is "satisfactory." Not noted is the finding that half of our students cannot write in even a "satisfactory"

manner as observed earlier.

The emphasis on averages in the majority report obscures the significant difference in scores assigned by the various readers. Some have stated that this spread indicates that the faculty cannot agree on what constitutes "good writing" but I think it merely indicates differences in the level of writing skills we either can or should expect of our freshman. These differences lead to the majority suggesting that the readers be trained if there is s the next round of review of the freshman writing programs I do not believe additional reviews are needed.

Further observations: It may be that we do not have either the resources or the will to improve these programs, i.e., to improve the writing skills of our students, but CEP and the campus administration should make this determination. It has been pointed out to me that Derek Bok's new book "Our Underachieving Colleges" "...reviews what empirical evidence on college writing programs exists and it is not a pretty picture." Perhaps UCSD is not alone in not teaching its students to write effectively. I appreciate as noted in the majority report that "... writing is a complex skill...." but the results of the present exercise establish once again that the existing programs at UCSD are not effective.

As an engineering faculty member for many years I have repeatedly heard my colleagues complain that our students, juniors and seniors, in preparing laboratory reports are unable to write clearly Furthermore employers of our students complain of the inability of our graduates to communicate satisfactorily. This hearsay evidence may apply to UCSD graduates in other disciplines. Accordingly if CEP wants further studies of the w writing skills of our students, an examination of those skills possessed by seniors could be of value.

June 21, 2007

TO: MARSHA CHANDLER,
Senior Vice-Chancellor Academic Affairs

FROM: Committee on Educational Policy

SUBJECT: College Writing Instruction Task Force Report

Over the course of many meetings this academic year, the Committee on Educational Policy (CEP) has considered both the majority and minority reports of the College Writing Instruction Task Force. The Task Force, as you know, consisted of five faculty members assigned to independently evaluate 348 freshmen essays randomly selected from the six colleges' core sequences and writing courses. The general conclusions of the Task Force were these: the overall quality of freshman writing is satisfactory; the quality of student writing varies across the six colleges; and, the quality of student writing does not improve significantly over the course of the first year of writing instruction. The minority report does not dispute the findings of the majority report, rather it disputes the report's "reticent tone" and its failure to state that "...the writing programs do not improve the writing skills of our freshman."

The Committee considered both of the Task Force's reports carefully. Additionally, we reviewed earlier reports on writing instruction and sought input from the Council of Provosts. The Committee accepts the conclusions of the Task Force majority report. **However, we are distressed by the fact that the quality of freshmen writing is, at best, satisfactory and that the report found no evidence that the college writing programs improve the quality of our students' writing.** This diminishes the level of instruction provided at all levels of undergraduate education. In the course of undergraduate program reviews, departments repeatedly report that they must modify their major curriculum on account of the poor writing skills of their students. There is a well-articulated frustration within the faculty regarding the lack of resources available to address these problems.

CEP is unwilling to accept the status quo regarding the quality of student writing. Furthermore, CEP is adamant that there be a renewed commitment to the success of writing instruction. The Committee continues to endorse the recommendations and basic principles of the 1998-1999 CEP report regarding writing instruction 1) College writing instruction at UCSD serves the entire campus; 2) College writing instruction can be expected to provide skills that are transferable to a variety of fields; 3) The central task of the college writing programs is to teach university-level writing.

That said, the Committee considered ways to address the concerns. We will recommend that the Senate establish a joint CEP/Graduate Council Committee during the 2007-2008 academic year charged to assess various models of writing instruction. Some of the issues this workgroup would review include:

1. Assess the use of graduate student teaching assistants in the college writing programs. This would include a careful review of the training offered to graduate students, performance

evaluations, and an assessment of how using graduate student TAs in the college writing programs impacts departments and graduate education (e.g. TA allocation formulas, graduate student recruitment). Additionally, the committee should assess what resources would be necessary to establish a common set of goals across colleges for training graduate TAs.

2. Review of the writing instruction models currently employed at UCSD, other UCs, and at other institutions with successful writing instruction programs (i.e., review of best practices).
3. Assess whether campus resources for writing instruction are currently used effectively or whether these resources could be more effectively deployed to achieve the University goal of generating a measurable improvement in undergraduate student writing during the freshman year.

In addition, the Committee is aware that there is some departmental interest in establishing writing courses that would serve in place of, or supplementary to, the college writing instruction programs. The Committee strongly recommends that the SVCAA entertain such experimental proposals. We envision that such proposals would have a strictly defined trial period of not less than 3 years, that students may have the option to “opt out” of their college writing programs to participate in these trials and that there would be continuous assessment and reporting to the CEP.

Scott Ashford, Chair
Committee on Educational Policy

cc: M. Appelbaum H. Powell D. Wulbert
J. Posakony B. Sawrey Chronfile

Horstmann, Bonnie

From: Hemingway, Gillian
Sent: Thursday, June 21, 2007 5:30 PM
To: Horstmann, Bonnie
Subject: FW: CEP Memo-- College Writing Instruction Task Force Reports

Attachments: CEP Recommendation to SVCAA.doc

From: Woolridge, Mary
Sent: Thursday, June 21, 2007 5:13 PM
To: MChandler
Cc: Appelbaum, Mark; 'bsawrey@ucsd.edu'; 'dwulbert@ucsd.edu'; 'jposakony@ucsd.edu'; 'hpowell@ucsd.edu'; Woolridge, Mary; 'sashford@ucsd.edu'
Subject: CEP Memo-- College Writing Instruction Task Force Reports

Dear SVC Chandler,
Attached is CEP's memo regarding the College Writing Instruction Task Force Report.
A signed copy of the memo will be delivered to you for your records.

Mary Woolridge
Academic Senate



CEP
Recommendation to SVCAA

MCC
C: COP C.F.
- Writing Prg. Dir. D.W.G.
SAN DIEGO: COUNCIL OF PROVOSTS -0106
CHAIR OF THE COUNCIL

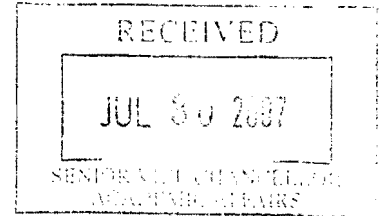
- BS (7/31 backlg)
- BAS
- BT

July 30, 2007

To: Marjorie Caserio
Acting Senior Vice Chancellor for Academic Affairs

From: Council of Provosts

Re: CEP Memo of June 21, 2007 concerning the College Writing Programs



The Council of Provosts has received a copy of CEP's memo of June 21, 2007 to Senior Vice Chancellor Chandler concerning the College Writing Instruction Task Force Report submitted to CEP in May 2006. There is one crucial point on which we agree with the CEP memo: "that effective writing instruction is of paramount importance in undergraduate general education." We enthusiastically endorse the application of the best writing pedagogy possible. However, we dispute the conclusions that CEP has drawn from the Task Force Report and oppose the course of action that CEP now proposes to the Academic Senate. The study, we feel, was at best inconclusive (by the review committee's own admission, there were many unanswered questions and methodological hurdles that were not solved). Informed colleagues have called the study seriously flawed in its structure and methodology (one of several examples: the structure and function of writing program essays are wholly different from the Subject A standard that was used to evaluate the essays). The current CEP proposal, made without any discussion or consultation with COP or the writing directors, strikes us as inimical to the process of shared governance and the educational interests of our students.

In his cover letter of June 22, 2006 that accompanied the Task Force Report, CEP Chair Curtis closed by stating, "We hope that a collaborative process of self-examination will take place, with next [2006-07] year's CEP as an active participant." COP, and subsequently, the writing program directors, responded to the initial report with observations, questions, and suggested next steps—including reactivation of the external review process initiated by Senior Vice Chancellor Chandler several years ago—but never received any communication back from CEP. Although the writing review report was not a standard departmental/program evaluation, we had expected that the findings would generate give-and-take between the reviewed programs and CEP, and further, that the provosts and writing directors would have the opportunity to address identified problems before CEP proceeded unilaterally.

Remarkably, CEP makes no mention of college participation in its June 21, 2007 memo to Senior Vice Chancellor Chandler and proposes the establishment of an Academic Senate Committee to review the writing programs. We vigorously support the continued examination of writing and the definition of best pedagogical practices. CEP has not identified the colleges as participants in the evaluation of their own educational programs. We continue to support an impartially conducted longitudinal study by an external panel to evaluate this critical area of "paramount importance to undergraduate education."

For these reasons, we hope that you will support the following:

- Reactivate the external review of the writing programs initiated in 2002 by SVCAA Chandler. To facilitate that process, writing program directors met this year and drew up the list of common goals that the 2002 panel of outside experts requested before the review could be completed. The importance of external scrutiny is reflected in the fact that other undergraduate programs are subject to periodic reviews which include one or more external reviewers; yet an external review of the writing programs has never been carried to conclusion. The reviews of the colleges' general education requirements which began last year with Revelle and which include at least one external reviewer do not focus on the writing component of those requirements (as the Revelle review demonstrates) nor is their brief to consider the college writing programs as such. An external panel of experts in writing pedagogy would provide the most informed and objective assessment of those programs as a group.
- Utilize the results of the external review whose reactivation is requested above to clearly define the objectives of any new pilot programs of the kind now being proposed by CEP, before funding for any such pilot program is approved.
- Upon completion of the external review, convene a joint Academic Affairs--Academic Senate Task Force to consider the results of the review and make any recommendations deemed necessary to ensure that writing is taught as effectively as possible across all four years of undergraduate study at UCSD.

We look forward to further discussion with you about the writing programs.

cc: Kim Griest, Chair CEP
Barbara Sawrey, Associate Vice Chancellor, Undergraduate Education

June 26 2006

Dear Chair Curtis and Other Colleagues:

I'm having some trouble with my UCSD email this morning so this is coming to you from my Columbia account.

I think the majority of members of the Writing Instruction Review Committee, not having met with writing program directors, TAs, or students, and not having reviewed the experience of other campuses or other writing program reviews, made the right decision to limit our report to the data we did review. Professor Libby criticizes the majority for its "reticence," but reticence is exactly what was called for.

If I had been asked to speak beyond the data, I would have drawn on my experience chairing the review of Muir Writing several years ago and my teaching in Marshall College's DOC sequence and, of course, my experience reading undergraduate term papers and exams at UCSD since 1981. And then I would have added the following:

Faculty and TAs in UCSD's writing programs work their tails off. UCSD invests many instructional dollars in these programs to insure small section size. Would the programs be better if section sizes were smaller still? Yes. Would they be better if TA's were better trained? No doubt. Would they be better if faculty and writing directors provided better prompts for student essays? Absolutely. I cringed as I read over the essays last summer to read one of the DOC prompts that was too complicated and very confusing. I cringed because I was among the faculty who wrote the damn thing! We don't always get it right. In my judgment, Revelle's program got it right more consistently than the other programs, and Warren got it wrong more often than the others. But none of this is as obvious and easy in practice as it seems at a distance.

How much do our efforts to teach writing to freshmen matter, in the end? None of us can answer that. We are plagued by optical illusions in judging how well students read and write. All we can recall from our own days as students is our OWN writing (although surely even this memory is distorted). We rarely saw the writing of our fellow students. The only thing we can be reasonably sure of is that we who are faculty at leading research universities got a lot of A's in college and our fellow students tended to get B's and C's. Our sense of the quality of writing in the good old days is thereby inevitably and invariably distorted upward.

Nonetheless, I am convinced (by casual observation and anecdote) that students do not read as much or as well as they used to. A 2002 NEH survey substantiates this. In 1982 60% of 18-24 year olds read at least one book of fiction, short stories, poetry, or plays in the past year; in 2002 the figure was 43%. TV had little or nothing to do with this decline. The Internet is not ruled out as a culprit. Newspaper reading among college students is vanishingly small. I'd be

interested to know how much our students use email -- they're speaking on their cell phones so often, when would they have time?

Does this make our students shallow? No. They have vastly more information about many more things than nearly any 40 or 50 or 60 year old did at 20. But they are novices at reading and writing, even with 4.0 GPA's from California high schools. Their high school textbooks are multi-media extravaganzas that depend little on the sustained prose we want them to learn to write. Their knowledge, even before the Internet, was coming to them in discrete bits, not in narrative prose.

Professor Libby implies that if the faculty complain loudly enough, we can design a freshman writing program that will overcome all of this, at least more than present programs do. Well, he may be right, but I think it would require a kind of Marshall Plan investment that neither UCSD nor any other research university is ever likely to make.

On the specific matter of whether the "improvement" finding is spurious: I brought this topic to the attention of the committee. It is not a decisive objection to the data but I think and most of my colleagues agreed that it raised a genuine question. The first of the four student papers we examined was NOT the first paper students turned in. It was the FINAL DRAFT of what the student turned in as a rough draft. The student then worked with peers and with TAs to turn the rough draft into a finished product. Any learning the student did in these several weeks of intensive writing instruction IS NOT CAPTURED BY THE DATA. Is it conceivable that students learned A LOT in this time? Is it even conceivable that students learned MOST of what they learned about meeting the demands of college writing instructors in this first encounter? Of course it is. Professor Libby is simply mistaken to say that the "improvement" finding is "incontrovertible."

Sincerely,

Michael Schudson



To: Charles Curtis, Chair, CEP
From: Alan Houston
Re: Freshman Writing Evaluation
17 July 2006

I write in response to your memo of 22 June, inviting comment on Paul Libby's minority report. I have read the replies of Michael Schudson and Sandy Lakoff, and am in general agreement with both. To them I would add the following:

Professor Libby's "main concern" is the "reticent tone" of the majority report. He believes that the majority has taken too narrow a view of its responsibilities, and wishes that we had done more to interpret and establish the significance of our findings. I believe these criticisms mistaken.

This committee was charged to assess freshman writing at UCSD. We asked whether the six college programs of writing instruction, as they currently exist, are successful in teaching students to meet the demands of college writing. With this goal in mind, we examined a large and representative sample of student portfolios. Each portfolio included four writing samples taken at different points during a single student's participation in a college writing program. Our findings are clear and clearly stated: freshman writing is, on average, satisfactory, though sometimes marginally so; there are statistically significant differences in the quality of student writing across the six colleges; and the quality of student writing did not improve as a result of writing instruction. These findings are striking, and raise a host of important questions regarding our students and the writing instruction they receive. Why don't students show more improvement of writing instruction? What explains variations in the quality of writing across the colleges? And, perhaps most insistently, can't we do better than this?

These are difficult questions, and our findings do not permit easy or complacent answers to them. No one should take comfort from the current state of undergraduate writing at UCSD. The return on investment in current programs of writing instruction appears to be small and unevenly distributed.

The work of this committee makes it possible, for the first time in UCSD's history, to test arguments and opinions concerning student writing and writing instruction against hard-won data. But this committee, as a committee, is not qualified to do more than provide that data. In accordance with our mandate, we concentrated on the labor-intensive task of evaluating student writing. We made no attempt to examine, individually or comparatively, the six college programs of writing instruction. We gathered no syllabi, interviewed no instructors, read no reports. We did not survey the experiences of other colleges and universities, nor did we meet to discuss what can or should be done at UCSD. This study is part of a much larger process of assessing the effectiveness of undergraduate writing instruction. To suggest that we *ought* to revamp current programs of writing instruction implies that we *can* do so in a way that will significantly improve undergraduate writing. This raises complex questions of pedagogy and policy. The appropriate venues within which to debate these issues are CEP, the Academic Senate and the six colleges. I sincerely hope that this will be done; our undergraduates deserve our best efforts.

August 14, 2006

PROFESSOR CHARLES CURTIS
Chair, Committee on Educational Policy

Re: Freshman Writing Evaluation

Thanks for your call to comment on the reports on freshman writing, as produced by the writing review committee and CEP.

I agree with CEP in emphasizing the importance of writing instruction at UCSD. In my view, most of UCSD's student writers (including graduate students) are not as good as they ought to be, and most entering freshmen write very poorly; nevertheless, great improvement often results from disciplined, rationally accessible writing instruction, conveyed with high expectations of students' willingness to learn. Few entering students have the skill in logic, the knowledge of words, the sensitivity to tone, and the habit of organization that make for good writing. Writing skills can, however, be taught.

Perhaps what I've said is all subjective opinion, incapable of support by statistical analysis. Its basis is merely a wide acquaintance with professional writing and many years of teaching at UCSD. Perhaps the most useful thing I can do at present is to mention some of the interesting mysteries of the reports. I will emphasize mysteries involving the Revelle College Humanities Program, because that's the one I know best.

1. Writing is taught in very different ways among the six writing programs, with Revelle using the most Aristotelian, traditional, and "old-fashioned" methods, and Warren using the most recent and "radical" ones. I don't think there's a respectable method of writing pedagogy currently employed in the United States that is not employed at UCSD. Why, then, according to the reports, do none of these methods produce statistically significant improvement in student writing? I refuse to believe that the answer is either, "Students can't learn," or "No one at UCSD knows how to teach."
2. Is there a relationship between the mystery in (1) and the fact that the distance between the earliest and the latest of the four evaluated writing assignments was (in respect to the Revelle program) a mere four months (c. 2/1/05 - 6/1/05)? (The report of the writing review committee, p. 2, inaccurately represents this as "two years.") Perhaps four months doesn't allow enough time for more improvement to register itself statistically.
3. Before producing the first writing assignment that was evaluated by the committee, Revelle students had already received four weeks of writing instruction, including most of what we consider the basics; they had also turned in a first, ungraded writing assignment and received extensive commentary on it. Does this fact have anything to do with the higher scores ("first paper" and "improvement") given to the writing of Revelle College students?
4. At Revelle, ERC, Marshall, and Sixth, writing instruction is integrated with instruction in a

sequence of courses in cultural and intellectual history. The first three of those colleges achieved the highest global writing scores. Does this suggest that “integrated” courses challenge and support student writing in ways that “stand alone” courses may not?

5. The report of the writing review committee, p. 2, wonders whether there may be a relationship between writing proficiency and the “uneven distribution across colleges of students apt by area of academic interest to have weaker skills in writing or, for that matter, apt to have more demanding courses to work on apart from their freshman writing course.” This suggestion seems plausible. Yet Revelle, with the highest “global” writing score, has very few humanities majors and a very demanding array of non-literary required courses--a combination that might be expected to give it the lowest writing score.

6. The individual graders’ ratings of writing improvement ranged from 1.87 (signifying that students wrote worse after instruction than before) to 2.66 (signifying that students improved about two-thirds of a grade after a few months of instruction). The mean of the ratings is 2.22 (about one-fifth of a grade—perhaps not enough to raise a C to a C+). Since there were only five reviewers and their ratings differed significantly, is the mean a statistically accurate measure of the students’ writing improvement?

Other mysteries may occur to me. In the meantime, I want to take this opportunity to request a meeting with CEP to discuss these issues. If I can be of help in any way, I hope you will let me know.

As always,

Stephen Cox
Professor, Literature
Director, Humanities Program

To the CEP

Re: The Report on the Writing Instruction Review

The committee did its best to follow instructions but the experience was frustrating because the mandate was narrow.

We were asked to grade for style a large number of student essays, carefully selected to provide a comparative view of the effectiveness of the various college writing programs. When our grades were processed statistically, they revealed some disparities in our evaluations, little evidence of improvement as a result of the writing courses, and some significant differences among the colleges that could not be accounted for by the background of the students. Grading subjectivity did not skew the findings of most members of the committee; only one of us was an “outlier.” To minimize grading disparity, the majority thought that if the exercise were to be repeated, a pre-grading prep session would be helpful. The majority report took note of the findings and cautioned that they may not be altogether conclusive.

I would add a few other considerations:

1. The quality of freshman writing was not as bad as I had expected it to be. The CEP should bear in mind, however, that the papers we saw were from the better-prepared students.

2. A major reason we found little improvement is that the quality of the early essays was not horrendous. The reason may be, as Schudson says, that the first paper we examined had been revised.

3. Unlike other members of the committee, I had no direct involvement in any of the college writing programs so I was not as sensitive as they to the differences in approach of the various colleges and could not comment on the comparative merits of different approaches.

4. The process was frustrating because we were asked to assign scores but were unable to learn what problems the writing instructors felt were most acute or make any suggestions for improvement to them, because we had no contact whatever with any of them. I took notes on common mistakes and weaknesses in the essays and would have called them to the instructors' attention if the opportunity had presented itself.

On a more personal note:

In over fifty years before the academic mast, I have put a fair amount of time and effort into improving writing – my own, my students', and my peers'. A piece I did for the LA Times last year musing about the great changes in education wrought by technology cited a favorite reference to the art of rewriting:

On my bulletin board is a clipping from the New York Times Book Review quoting the humorist S.J. Perelman, who was once asked how many drafts he did of a story: "Thirty-seven," he replied. "I once tried 33 but something was lacking, a certain, how shall I say, *je ne sais quoi*. On another occasion I tried 42 versions, but the final effect was too lapidary."

Perelman's pearl of wisdom resonates with my own experience. I first realized how much of teaching involved writing instruction as a Teaching Fellow and then a fledgling faculty member at Harvard. In addition to directing discussion sections of large classes, I taught in the tutorial program, going over students' essays and showing them how to improve their writing. Sometimes, I would sit the student down next to me at the typewriter and bat out revised versions of a few paragraphs of his report on what he had read. In the junior and senior tutorials, we had weekly sessions with individual students in which we worked with them on rewriting. (After a few years, I was put in charge of the program, as Head Tutor in the Department of Government.) One of my first tutees, who went on to get a PhD and became a professor, sent this e-mail the other day, after he had taken part in a class reunion program at Harvard:

Well, my 50th reunion was great fun. The highlight for me was that my symposium presentation to my classmates, the first afternoon, went over extremely well...So, I may not be rich, but you did a wonderful job of mentoring a B minus student who cut a lot of classes during his undergraduate years. Thank you, thank you.

When we founded the Political Science department at UCSD, we did what we could to replicate the Harvard tutorial/Princeton preceptorial program by creating a senior honors program. It works very well and attracts about twenty of our majors each year. Last year, I agreed to take on — for one last time — a senior writing an honors thesis. I am pleased to say he won the department prize and a national honor as well.

We've all had gratifying experiences like these with both undergraduates and graduate students because teaching writing is a large part of teaching, especially in the humanities and social sciences. In my case, it has also extended to my work as a referee. In countless cases, I have made detailed suggestions for revision to publishers, journals, and colleagues who have sent draft manuscripts.

But all this experience has taught me that you can't teach writing en masse. It can only be done one-on-one, or one manuscript at a time. And of course, you only really perfect your own writing by working at it constantly and paying attention to good writing when you encounter it.

I do believe that clear, effective writing can be "taught" and learned, even though some of us are more teachable than others when it comes to writing as it does to other skills. Teachers can and should give students pointers and incentives and review examples with them. I always remind students, for example, that every essay and every paragraph should start with a topic sentence and that paragraphs must follow logically

and rhetorically from each other. In my courses, students are given the option of writing a term paper under a “double deadline” system to encourage revision. It’s more work for the TAs and me but it’s well worth it. To cite from the syllabus of one course:

Term paper is optional. Term paper grade will be counted only if it raises the course grade. Course grade would then be composed of midterm (20 percent), paper (40 percent) and final exam (40 percent). Term paper must be on a recommended or approved topic (see list below), between 2,000 and 2,500 words. Final deadline for submission of term paper is Nov. 29. If submitted **by no later than Nov. 15**, it will be returned with comments and a provisional grade one week later, and may be revised and resubmitted by the final deadline; the higher of the two grades will count.

To encourage good work, I always ask the student who has written the best answer on the midterm to read it aloud in class. (Invariably, the student who is asked is mortified and reluctant but after the answer is read there is spontaneous applause.) When time permits, I also have an outstanding term paper read in class. And I have recommended a number of the papers for presentation in the annual undergraduate research symposium – an excellent program that means a lot to students after they graduate.

In addition, to combat students’ abysmal ignorance of the world and lack of background (in this instance in the politics of the Middle East), I now require that they subscribe to the weekday edition of the New York Times for the quarter. I have an ulterior motive, to be sure. I want them to keep up with news from the Middle East but I am also hoping they will discover how to write good clean prose, how to present an argument (from the op eds, columns, and letters) and how interesting a great newspaper can be.

These are techniques that can readily be used in many courses, without imposing great burdens on faculty or TAs, but they are not enough to overcome the handicaps in speech and writing that are so pronounced among our undergraduates. (You only have to walk around the campus and listen to snatches of conversation to realize how impoverished they are in the use of language: “So I’m like...and she goes... and I’m like, and I go,” etc., etc...) What else could be done?

First of all, UCSD needs to decide whether it wants to be serious about undergraduate education or whether it wants to concentrate on being a great research university, allowing undergraduates to learn by osmosis. And let me be specific. A university seriously concerned with undergraduate education would never have allowed Jonathan Saville to take early retirement. How many times does somebody have to be named the best teacher at Revelle for the university to recognize it has a great teacher? (I think he won the honor four times.) He should have been appointed the William Strunk and E.B. White Professor of Prosody (or maybe the Perelman Professor of 37 Rewrites) and put in overall charge of the college writing programs. (While Jonathan was still doing music reviews for The San Diego Reader, I used to tell students that if they wanted to learn how to write, they should simply pick up the Reader for free each week.)

And while I'm grouching, let me also note that the Engineering School should have taken my advice and hired a friend of mine who ran an outstanding humanities for engineering students program at the University of Colorado – instead of thanking me politely and ignoring the suggestion.

Unless we take this mission more seriously, we will never become as good overall as the Ivy League universities or Oxbridge. They are not just great research universities; they are also places that take undergraduate education seriously and therefore attract the best students (including some whose parents teach at UCSD!). The proof of the pudding is that while we turn out lots of students who do well in graduate and professional schools, especially in the sciences, engineering, and medicine, we never produce any who win Rhodes, Marshall, or the other fellowships that stress well-roundedness and articulateness. We are handicapped, to be sure, because we don't have lavish graduate fellowships that supply us with lots of cheap labor: at Harvard, much of the tutorial burden was borne by graduate students. But we could do a much better job of improving student writing if we made it part of a general effort to upgrade undergraduate education and experimented with some innovative efforts to put faculty, emeriti, and non-faculty part-timers to work helping students learn to write, not just as freshmen but all through the undergraduate program. Maybe the next committee the CEP appoints on this issue should be given a broader mandate, not just to assess the value of the existing first year writing programs, but to think imaginatively about how they could be revamped and extended, so as to strengthen our approach to liberal education despite the financial constraints.

Sandy Lakoff