Writing Project

Variables of Class Level, College, Gender, and Trait. A single omnibus mixed model ANOVA was done to compare the effects of the between-subjects variables of Class Level, College, and Gender and the within subjects variable of Trait. Only the main effects of Class Level, $F(3, 226) = 4.21, \text{MSE} = .354, p < .01$, and Trait, $F(4, 904) = 17.08, \text{MSE} = .189, p < .01$, were significant (all other $F$'s $\leq 1.14$). Follow-up tests using Tukey’s HSD showed that freshmen scored significantly lower than sophomores, juniors, and seniors ($p$'s $\leq .05$). No other significant differences were found for class level.

Separate ANOVAs for the planned comparisons for each trait as a function of class level showed a significant effect of class level for every trait. For Purpose, $F(3, 266) = 4.32, \text{MSE} = .626, p = .01$, the follow-up comparisons using Tukey’s HSD (used for all reported follow-up comparisons reported here) showed that freshmen scored significantly lower than both sophomores and seniors. No other class level comparisons were significant. For Synthesis, $F(3, 266) = 4.41, \text{MSE} = .501, p < .01$, freshmen scored lower than both juniors and seniors. For Style, $F(3, 266) = 2.87, \text{MSE} = .456, p < .04$, only the difference between seniors and freshmen was significant, with freshmen scoring lower. Finally, for both Support, $F(3, 266) = 10.34, \text{MSE} = .499, p < .01$, and Mechanics, $F(3, 266) = 6.60, \text{MSE} = .447, p < .01$, follow-up comparisons showed that freshmen scored significantly lower than sophomores, juniors, and seniors, with no significant differences among these groups.

English 134. In order to test the hypothesis that post-test scores would be significantly higher than pre-test scores, a paired sample $t$-test was used. The result was significant, $t(55) = 7.20$. Post-test scores ($M = 3.01$) were significantly higher than pre-test scores ($M = 2.20$). As a follow-up, both the pre-test and post-test scores were compared to a constant of 3 (indicating average attainment on the holistic rubric) in separate one-sample $t$-tests. Pre-test scores were significantly lower than the criterion score of 3, $t(55) = -9.77, p < .05$. In contrast, post-test scores did not differ significantly from the criterion score ($p = .86$).

Pre-test and post-test scores were both correlated with final grades. Pre-test scores were weakly correlated with final grades, $r(54) = .24, p = .08$, whereas post-test scores were significantly correlated with final grades, $r(54) = .33$.

Oral Communication

Separate chi-square analyses confirmed that the observed frequencies for both the gender distribution and the college distribution did not differ significantly from the expected frequencies.

Differences in the mean scores for traits were analyzed in two separate repeated measures ANOVAs—one looking at all seven traits for the 75 students who had scores on all of the traits, and the second for the 102 students who had complete data on the six traits excluding the Use of Visual Aids. Both analyses were significant, $F(6, 444) = 5.70, \text{MSE} = .430, p < .01$ for the seven-trait comparison; $F(5, 505) = 6.62, \text{MSE} = .383, p < .01$ for the six-trait comparison. Follow-up pairwise comparisons using a Bonferroni adjustment showed the same basic pattern in both sets of analyses. Students’ trait scores were significantly higher for Language Use and Use of Supporting Materials than for Verbal and Non-Verbal Delivery scores. In addition, the presence of a Central Message was significantly higher than the Verbal Delivery score in both analyses; in the seven-trait analysis the presence of a Central Message was also significantly higher than the Non-Verbal Delivery score (all $p$'s $\leq .05$). There were no other significant differences.
Diversity Learning

DLO1. An omnibus factorial ANOVA was done to analyze the total mean scores for DLO 1 as a function of Class Level, College, Survey Mode (in-class, online), and Sex. Table D2 shows the breakdown of scores by various student categories. It should be noted that the cell sizes were too small or had no observations to support looking at possible three-way and four-way interactions.

The ANOVA results for DLO 1 showed significant main effects for Class Level, $F(2, 157) = 10.02, MSE = .451, p < .01$; College, $F(5, 157) = 5.88, MSE = .451, p < .01$; Survey Mode, $F(1, 157) = 18.40, MSE = .451, p < .01$; and Gender, $F(1, 157) = 6.84, MSE = .451, p < .01$. Significantly higher scores were evident for the online survey and for males. Follow-up Tukey HSD tests on the main effect of Class Level yielded evidence for value added: both seniors and juniors scored higher than freshmen but did not differ from one another. With regard to the main effect of College, the follow-up Tukey HSD tests showed that CAFES students scored significantly lower than OCOB, COSAM, and CENG students. No other College differences were significant.

There was also a significant interaction of Sex by Class Level, $F(2, 157) = 4.89, MSE = .451, p < .01$. The value added was more apparent in men, such that male seniors had significantly higher scores than male freshmen: $M($freshman$) = 1.66, M($juniors$) = 2.00, M($seniors$) = 2.48$. This was not so with women: $M($freshman$) = 1.48, M($juniors$) = 1.83, M($seniors$) = 1.71$, whose scores did not differ as a function of Class Level. It should be noted that marginally significant interactions were also present for College by Level ($F = 1.91, p = .05$) and College by Survey Mode ($F = 2.36, p = .06$) but are not broken down further here because of concerns with sample sizes.

DLO 2. As with DLO 1, an omnibus factorial ANOVA was done to analyze the total mean score for DLO 2 as a function of Class Level (freshman, junior, senior), College (CAED, CAFES, CENG, CLAM, CSM, OCOB), and Survey Mode (in-class, online). Sex was not included in the analysis. Table D3 shows the breakdown of the scores by various student categories.

There were significant main effects of Class Level, $F(2, 206) = 8.53, MSE = .529, p < .01$; College, $F(5, 206) = 4.64, MSE = .529, p < .01$; and Survey Mode, $F(1, 206) = 4.11, MSE = .529, p < .05$. Again, the online survey mode resulted in significantly higher scores.

The Class Level effect showed that while there were no differences between junior and senior scores, both seniors and juniors scored significantly higher than freshmen, with no differences between their scores. The College Level effect showed that COSAM students scored significantly higher than CAFES and CENG students, with no other College Level differences reaching significance. There was, however, a significant interaction between Class Level and College, $F(10, 206) = 1.92, MSE = .529, p < .05$, such that among freshmen, COSAM students ($M = 1.75$) scored significantly higher than OCOB students ($M = .90$) and that among seniors, COSAM students ($M = 2.49$) scored significantly higher than CENG students ($M = 1.45$). Especially with regard to the college-level results, including these interactions, small, unequal sample sizes mean that one should not over interpret these results.

DLO 3. Table D4 presents the mean scores for DLO 3. The omnibus ANOVA using Class Level, College, Survey Mode, and Gender yielded three significant main effects and no interactions. Class Level, $F(2, 179) = 23.18, MSE = .564, p < .01$; College, $F(5, 179) = 4.40, MSE = .564, p < .01$; and Gender, $F(1, 170) = 12.01, MSE = .564, p < .01$, were all reliable.

USCP. Because fulfillment of the USCP requirement is the major curricular path for developing diversity-related competence, a separate t-test was done to compare mean scores for juniors and seniors.
only collapsed across both Class Level and DLO. Although the overall average score for juniors and seniors who had not completed a USCP course (n = 63, M = 2.02) is lower than the score for juniors and seniors who had completed a USCP course (n = 205, M = 2.18), this difference was not statistically significant.

**Service Learning.** The overall average score for juniors and seniors who had not completed a service-learning course (n = 155, M = 2.08) was lower than the score for juniors and seniors who had completed a service-learning course (n = 137, M = 2.19), but again a t-test showed that this difference was not statistically significant. The percentage of student essays with scores in the “3 = moderate” or “4 = complex” levels, was 32% for juniors and seniors who had not completed a service-learning course and 40% for juniors and seniors who had completed a service-learning course. Similar to USCP, these assessment results do not indicate that service learning makes a large positive contribution to diversity learning as defined by the DLOs.

**Lifelong Learning**

Table LL1 presents the mean scores in terms of percent correct for five questions for which there was a single response, i.e., check one, that could then be coded as correct or incorrect. Five 2x2 ANOVAs were done to examine whether the correct response on each item was related to Class Level and Instruction. It was hypothesized that both factors would be related to success on the questionnaire, with upper-division students doing better than lower-division students and students who had received library instruction doing better than students who had not received instruction. Across all analyses, no significant interactions were present (all \( F \)'s ≤ 1.04).

Significant main effects of Class Level (value added) were found for three of the five items; in all cases, upper-division students did better than lower-division students. Significant effects of Instruction were found for the correct identification of the thesis statement and the citation example. For thesis statement/promising research question, upper-division students, \( M = .783 \), did significantly better than lower-division students, \( M = .676 \), \( F(1,794) = 12.02, \text{MSE} = .187 \), and students with library instruction, \( M = .774 \), did significantly better than students who had received no instruction, \( M = .771 \), \( F(1,794) = 4.29, \text{MSE} = .187 \). In the correct identification of the citation example, again upper-division students, \( M = .513 \), did significantly better than lower-division students, \( M = .384 \), \( F(1,757) = 14.53, \text{MSE} = .137 \), as did students with library instruction, \( M = .501 \), compared to students without library instruction, \( M = .429 \), \( F(1,757) = 6.81, \text{MSE} = .137 \). Upper-division students, \( M = .784 \), did significantly better than lower-division students, \( M = .680 \), on the correct selection of the search term that would yield the fewest results, \( F(1,782) = 10.57, \text{MSE} = .187 \). Finally, there was a marginal effect of class level on the correct selection of the search term that would yield the most results, \( F(1,782) = 3.08, \text{MSE} = .240, p = .08 \); upper-division students (\( M = .623 \)) were slightly more apt to identify the item correctly as compared to lower-division students (\( M = .556 \)). The question on the ethical use of ideas showed no significant effects of either Class Level or Instruction (\( F \)'s < 1.00).

**Ethics**

Out of 31 points possible, the average exam score was 12.45, i.e., students answered 40% of the questions correctly. Because of small and uneven sample sizes and concerns regarding the distributions of the data, separate Kruskal-Wallis tests were run to compare the total scores as a function of Class Year (see Table E2) and College (see Table E3); the result for Class Year was not significant (\( F < 1.50 \)). There was no evidence of value added on the ethics scores, though this may have been a function of small sample sizes. The visual pattern of the data when comparing first-year students to fourth- and fifth-year students is in the predicted direction.
The result for College was significant, $\chi^2 (5, N = 264) = 14.95$, $p = .011$. Separate Mann-Whitney U tests as follow-ups showed that students in the College of Science and Math scored significantly higher than students in all the other colleges. No other differences among colleges were significant.

A mixed model analysis of variance with one between subjects factor (Course) and one within subjects factor (Trait) compared the four different traits as function of course enrollment. There were no effects involving having taken an ethics course (all $F$s $\leq 1.75$). The repeated measures factor of Trait showed a significant difference among the scores, $F(3, 786) = 13.465$, $MSE = .032$, $p < .01$. Follow-up pairwise comparisons using a Bonferroni adjustment for multiple comparisons showed that students scored significantly higher on “Application of Ethical Theories/Concepts” as compared to both “Understanding Different Ethical Theories/Concepts.” and “Ethical Issue Recognition.” Students also did significantly better on the “Evaluation of Different Ethical Perspectives/Concepts” as compared to their “Understanding Different Ethical Theories/Concepts” (all $p's < .01$). Finally, there was a marginal effect such that “Ethical Issue Recognition” was slightly better than “Understanding Different Ethical Theories/Concepts” ($p = .06$). No other comparisons were significant.