



**2011-12 Assessment of General Education Outcomes**  
*Written Communication*

**Office of Institutional Assessment and Evaluation**

**September, 2012**

*“Working together to create a culture of evidence-based decision making...”*



**Office of Institutional Assessment and Evaluation**  
*“Working together to create a culture of evidence-based decision making...”*

**2011-12 Assessment of General Education Outcomes**  
***Written Communication***

### **The Assessment of Written Communication**

In Spring Semester 2012, the Office of Institutional Assessment and Evaluation collected data on student performance in English 003 from the English Department for the purpose of determining the extent to which students demonstrated competence in the general education outcome, *written communication*.

During the semester, students were required to compose essays in the areas of Style Analysis, Definition, Investigative, and Solution. A descriptive analytic rubric was used for grading each essay based on five criteria: (1) Organization, (2) Development A, (3) Development B, (4) Style, and (5) Grammar/Mechanics. A copy of the analytic rubric is in the Appendix.

For each criterion, a score is given in one of six levels, with maximum of 20 points. The top level is *Strong*, which has a score range of 18-20 points. The second level is *Satisfactory*, which has a score range of 16-17 points. The third level is *Passing/Acceptable*, which has a score range of 14-15 points. A score that is lower than 14 is considered a failing grade. Failing scores fall into three levels: *Failing, Inadequate* (12-13 pts), *Failing, Seriously Flawed* (1-11 pts), and *Total Failure* (0 pts). A student’s total score on an essay is the sum of all criterion scores. The total score on an essay can range from 0 to 100.

### **Description of Criteria on the Analytic Rubric**

The descriptions of each of the five criteria are as follows:

- **Organization:** Organizes ideas and connects them with clear transitions.
- **Development A:** Identifies and analyzes important features of the argument and supports the main points of the argument.
- **Development B:** Explains opposition and refutes it; uses documentation.
- **Style:** Control of language, including diction and syntactic variety.
- **Grammar/Mechanics:** Facility with conventions (grammar, usage, and mechanics).

For each criterion, a score is earned on one of six levels. The maximum number of points for each criterion is 20. The first/highest level is *Strong*, which has a score range of 18-20 points. The second level is *Satisfactory*, which has a score range of 16-17 points. The third level is *Passing/Acceptable*, which has a score range of 14-15 points. A score that is lower than 14 is considered a failing score. Failing scores fall into three levels: *Failing, Inadequate* (12-13 pts), *Failing, Seriously Flawed* (1-11 pts), and *Total Failure* (0 pts). A student’s total score on an essay is the sum of all criterion scores. The total score on an essay can range from 0 to 100.

## Studies on the Development of Students' Writing Skills in English 003

Scoring rubrics have become widely used in educational settings as a tool for assessing student writing skills (Andrade, 2000). When descriptive analytic rubrics are used, students receive a score on their performance on each writing criterion as well as a description of what that score means. The results can be used to make improvement in areas of weakness. Rubrics can also be used for purposes of evaluation and accountability. The analyses of students' performances across multiple sections of a course can inform the attainment of institutional level goals and objectives.

For Spring Semester 2012, two studies were conducted on the English 003 essay data. Study 1 focused on the effects of writing revisions of essays on the improvement of students' writing skills. Study 2 focused on the effects of writing multiple essays, without revisions, on the improvement of students' writing skills.

### Study 1

The literature on formative assessment provides guidance on the characteristics of effective assessment and feedback (Sadler, 1989; Hattie and Timperley, 2007; Hughes, 2010). Two important components of formative assessment and feedback include: (1) articulated *clear criteria* for assessing writing, and (2) provide *opportunities for improvement through revision* (Andrade, 2000).

*Procedure and Results:* During the course of the semester, students in one section of English 003 were assigned a research essay and submitted subsequently two revisions. Scores are given based on a descriptive analytic rubric. Table 1 presents the means and standard deviations of criterion scores for the three essays of original, first, and second revision. The valid "N" consists of the number of students who participated in this exercise. The average total score increased from  $M = 86.86$ , to  $M = 88.67$  to  $M = 89.08$ . In general, revisions have increase mean score for each criterion. However, the mean score for *Grammar/Mechanics* lowers after the second revision compared with the first revision. Variability of students' scores increases on *Organization* on both revisions, increases once on *Grammar/Mechanics* on the second revision, and decrease on all other criterion. Mean scores for *Grammar/Mechanics* are much lower than any other criterion. See graph presentation in *Study 1 Graphs* section at the end of the report.

**Table 1. Descriptive Statistics of Original Essay and Its Revision by Grading Criteria**

| Criterion                | Valid N | Mean           |                |                 | Std. Dev.      |                |                 |
|--------------------------|---------|----------------|----------------|-----------------|----------------|----------------|-----------------|
|                          |         | Original Essay | First Revision | Second Revision | Original Essay | First Revision | Second Revision |
| <b>Organization</b>      | 36      | 18.64          | 18.86          | 18.94           | 0.68           | 0.80           | 0.83            |
| <b>Development A</b>     | 36      | 18.25          | 18.56          | 18.72           | 0.84           | 0.70           | 0.70            |
| <b>Development B</b>     | 36      | 17.83          | 17.75          | 18.33           | 1.13           | 0.84           | 0.83            |
| <b>Style</b>             | 36      | 17.42          | 17.83          | 17.83           | 0.94           | 0.51           | 0.78            |
| <b>Grammar/Mechanics</b> | 36      | 15.03          | 15.69          | 15.22           | 0.51           | 0.47           | 0.54            |
| <b>Total Score</b>       | 36      | 86.86          | 88.67          | 89.08           | 3.98           | 2.54           | 3.20            |

Mean scores increase reflects an average improvement for the class. The progress of student writing performance through this revision exercise can be seen more clearly by individual student score. The following histograms show how total score improvement through revision. There are 8 students scored between 75 and 84 after submitting the original essay. After first revision, only two students remain in this score range and six more students increase their scores to the range of 85 to 89. The number of students in the range of 90 to 94 remains the same. After second revision, the number of student who scores in the range of 90 to 94 increase from 15 to 22, a net increase of 7, or a 47% increase. The number of students in the range of 85 to 89 reduced from 19 to 13. Only one student actually lowers his/her score after second revision. Larger graphs are listed in the *Study 1 Graphs* section at the end of the report.

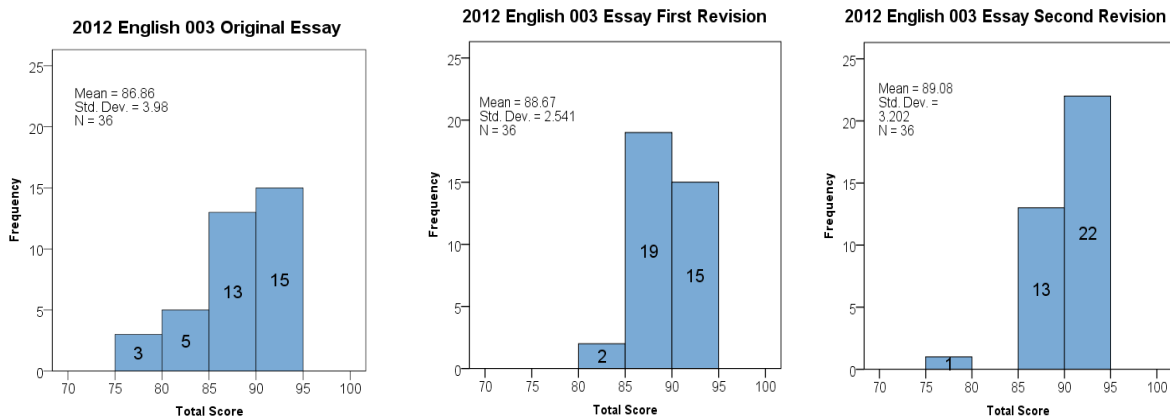


Figure 1.

## Study 2

Table 2 presents the descriptive statistics by each criterion for Study 2 in which students wrote essays without submitting revisions. There were a total of 187 valid observations. The scores of the original essay used in Study 1 are also included in Study 2, but scores of the revision essays from Study 1 are excluded. The Minimum and maximum scores for each criterion are 0 and 20, respectively. A mean score for each criterion was calculated across all valid observations. *Grammar/Mechanics* has the lowest mean score,  $M=15.06$ ,  $SD=1.57$  and *Organization* has the highest mean score,  $M=17.26$ ,  $SD=2.18$ .

**Table 2. Descriptive Statistics for English 003 by Grading Criteria - Spring Semester 2012**

| Criterion         | Valid N | Minimum | Maximum | Mean  | Std. Dev. |
|-------------------|---------|---------|---------|-------|-----------|
| Organization      | 187     | 10      | 20      | 17.26 | 2.18      |
| Development A     | 187     | 10      | 20      | 17.03 | 2.26      |
| Development B     | 187     | 0       | 20      | 16.04 | 3.23      |
| Style             | 187     | 10      | 20      | 16.53 | 1.95      |
| Grammar/Mechanics | 187     | 10      | 20      | 15.06 | 1.57      |
| Total Score       | 187     | 42      | 99      | 81.93 | 9.37      |

The bivariate correlation coefficients are shown in Table 3. All correlation coefficients are positive and significant at the 0.01 level. The correlation between *Grammar/Mechanics* and *Development B* is the smallest with a value of .334. The correlation between *Organization* and *Total Score* is the largest with a value of 0.916.

**Table 3. Pearson Correlations for Five Criteria and Total Score**

|                   | Organization | Development A | Development B | Style    | Grammar/Mechanics | Total Score |
|-------------------|--------------|---------------|---------------|----------|-------------------|-------------|
| Organization      | 1            | 0.833 **      | 0.652 **      | 0.866 ** | 0.530 **          | 0.916 **    |
| Development A     | 0.833 **     | 1             | 0.621 **      | 0.743 ** | 0.415 **          | 0.866 **    |
| Development B     | 0.652 **     | 0.621 **      | 1             | 0.602**  | 0.334**           | 0.823 **    |
| Style             | 0.866**      | 0.743 **      | 0.602 **      | 1        | 0.723**           | 0.911 **    |
| Grammar/Mechanics | 0.530 **     | 0.415 **      | 0.334**       | 0.723**  | 1                 | 0.650 **    |
| Total Score       | 0.916 **     | 0.866 **      | 0.823 **      | 0.911 ** | 0.650**           | 1           |

\*\* Correlation is significant at the 0.01 level (2-tailed).

### ***F-test and Post Hoc Analysis***

An Analysis of Variance (ANOVA) was performed on the five criterion means to determine if students performed equally across all criteria. The *F*-test shows that the means for five criteria are *not* equal and the overall difference is significant,  $F(4,930)=27.22, p<.001$  The results of *F*-test are shown in Table 4.

**Table 4. ANOVA for Five Criteria Used in the Grading Rubric**

|                | Sum of Squares | df  | Mean Square | F      | Sig.  |
|----------------|----------------|-----|-------------|--------|-------|
| Between Groups | 576.667        | 4   | 144.167     | 27.224 | <.001 |
| Within Groups  | 4924.952       | 930 | 5.296       |        |       |
| Total          | 5501.619       | 934 |             |        |       |

In order to compare the mean difference between criteria, a Post Hoc analysis was performed. From Table 2, we know that the variances are different. Therefore, equal variances are not assumed. Tamhane's T2 Post Hoc test was used. The results are shown in Table 5.

In Table 5, the mean differences are shown in the third column. The mean for *Organization* is significantly larger than any other criterion except *Development A*. The mean for *Grammar/Mechanics* is significantly smaller than any other criterion. The ranking for the mean is *Organization* >= *Development A* >= *Style* >= *Development B* > *Grammar/Mechanics*. The mean of

*Organization* is slightly larger than the mean of *Development A*, but the difference is not significant ( $p=0.984$ ). The largest mean difference is between *Organization* and *Grammar/Mechanics*, and the difference is 2.198.

**Table 5. Post Hoc Multiple Comparisons**

| (I) Criterion            | (J) Criterion     | Mean Difference (I-J) | Std. Error | Sig.  | 95% Confidence Interval |             |
|--------------------------|-------------------|-----------------------|------------|-------|-------------------------|-------------|
|                          |                   |                       |            |       | Lower Bound             | Upper Bound |
| <b>Organization</b>      | Development A     | .219                  | .229       | .984  | -.42                    | .86         |
|                          | Development B     | 1.214*                | .285       | <.001 | .41                     | 2.02        |
|                          | Style             | .722*                 | .214       | .008  | .21                     | 1.32        |
|                          | Grammar/Mechanics | 2.198*                | .196       | <.001 | 1.65                    | 2.75        |
| <b>Development A</b>     | Organization      | -.219                 | .229       | .984  | -.86                    | .42         |
|                          | Development B     | .995                  | .288       | .006  | .18                     | 1.81        |
|                          | Style             | .503                  | .218       | .196  | -.11                    | 1.12        |
|                          | Grammar/Mechanics | 1.979*                | .200       | <.001 | 1.41                    | 2.54        |
| <b>Development B</b>     | Organization      | -1.214*               | .285       | <.001 | -2.02                   | -.41        |
|                          | Development A     | -.955*                | .288       | .006  | -1.81                   | -.18        |
|                          | Style             | -.492                 | .276       | .545  | -1.27                   | .29         |
|                          | Grammar/Mechanics | .984*                 | .262       | .002  | .24                     | 1.72        |
| <b>Style</b>             | Organization      | -.722*                | .214       | .008  | -1.32                   | -.12        |
|                          | Development A     | -.503                 | .218       | .196  | -1.12                   | .11         |
|                          | Development B     | .492                  | .276       | .545  | -.29                    | 1.27        |
|                          | Grammar/Mechanics | 1.476*                | .183       | <.001 | .96                     | 1.99        |
| <b>Grammar/Mechanics</b> | Organization      | -2.198*               | .196       | <.001 | -2.75                   | -1.65       |
|                          | Development A     | -1.979*               | .200       | <.001 | -2.54                   | -1.41       |
|                          | Development B     | -.984*                | .262       | .002  | -1.72                   | -.24        |
|                          | Style             | -1.476*               | .183       | <.001 | -1.99                   | -.96        |

\*. The mean difference is significant at the 0.05 level.

Note: "Tamhane's T2" Post Hoc test is used, and equal variances are not assumed.

### ***Passing Rate by Criterion***

The passing competency requires a score of 70% or higher, i.e., 14 out of total 20 points. The results are shown in Table 6. The criterion with the highest passing rate is the *Organization*, which

has a passing rate of 90.4%. *Development B* has the lowest passing rate at 84.5%. The ranking for passing rate is *Organization* > *Style* > *Development A* > *Grammar/Mechanics* > *Development B*.

**Table 6. Passing Rate by Criterion**

| Criteria                 | Status | Frequency | Percent |
|--------------------------|--------|-----------|---------|
| <b>Organization</b>      | Fail   | 18        | 9.6%    |
|                          | Pass   | 169       | 90.4%   |
|                          | Total  | 187       | 100.0%  |
| <b>Development A</b>     | Fail   | 22        | 11.8%   |
|                          | Pass   | 165       | 88.2%   |
|                          | Total  | 187       | 100.0%  |
| <b>Development B</b>     | Fail   | 29        | 15.5%   |
|                          | Pass   | 158       | 84.5%   |
|                          | Total  | 187       | 100.0%  |
| <b>Style</b>             | Fail   | 21        | 11.2%   |
|                          | Pass   | 166       | 88.8%   |
|                          | Total  | 187       | 100.0%  |
| <b>Grammar/Mechanics</b> | Fail   | 24        | 12.8%   |
|                          | Pass   | 163       | 87.2%   |
|                          | Total  | 187       | 100.0%  |
| <b>Total Score</b>       | Fail   | 25        | 13.4%   |
|                          | Pass   | 162       | 86.6%   |
|                          | Total  | 187       | 100.0%  |

*Note:* Based on descriptive analytic rubric, the passing grade requires a score of 70% of the total score, i.e., 14 out of total 20 points.

## Summary

The ultimate goal for outcomes assessment is to use the results of assessment to improve student learning and performance. This report embraces this goal by emphasizing student performance at the criterion level. The disaggregation of scores at the criterion level vis-à-vis total scores at the essay level provides an advantage in implementing strategies for improvement. First, it informs the instructors in the Department of English about the strengths and weaknesses of students enrolled in English 003 in the aggregate. Second, student performance data at the criterion level provide instructors with specific targets for improvement in the teaching and learning

process. Third, analysis at the criterion level provides information about students' relative strengths and weaknesses among different types of essay composition (e.g., Style Analysis, Definition, Investigative and Solution). This information can help instructors balance time and effort devoted to helping students develop competency in each criteria and for each essay type.

For Spring Semester 2012, students enrolled in English 003 met the minimum performance criterion of 70% in *all* criterion areas. For all five criterion areas, students' performance exceeded the minimum performance criterion ranging from 14.5% to 20.4%.

It is also instructive to note that when students were allowed to revise their essays and were given feedback on their performance, their scores increased significantly and the variability in their performance decreased.

Similar to the results from Spring Semester 2010, and Spring Semester 2011, *Development B* and *Grammar/Mechanics* have the lowest mean scores (See Tables 2) and passing rates (Table 6). Enhanced training in these two areas can effectively improve students' writing competencies in English 003. The results for Spring Semester 2012 exceeded the results for Spring Semester 2010 and 2011 for every criterion except *Grammar/Mechanics*. The Office of Institutional Assessment and Evaluation (OIAE) will continue to monitor student learning outcomes in written communications to identify discernible trends.

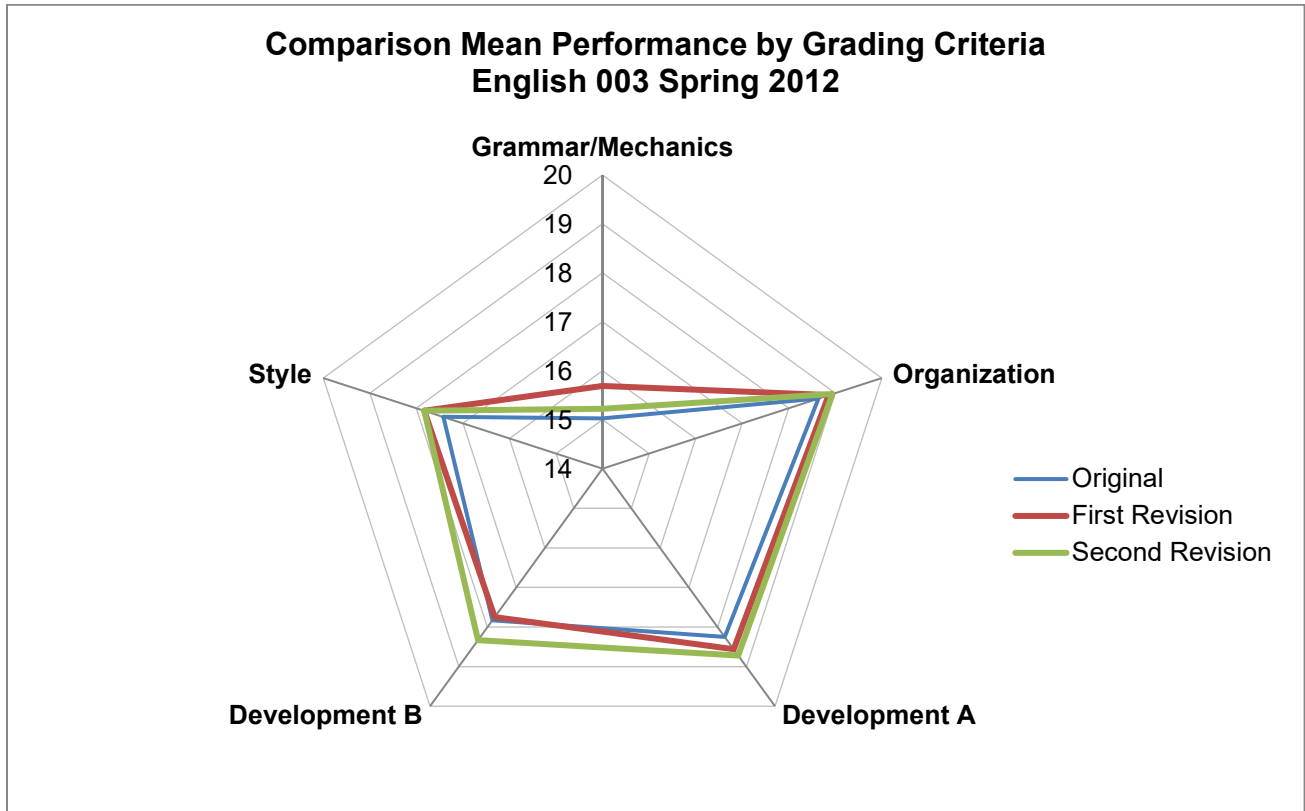
## References

- Andrade, H. (2000). Using rubrics to promote thinking and learning. *Educational Leadership*, 57(5), 13-18.
- Hattie, J. A., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112.
- Hughes, G. (2010). Formative assessment practices that maximize learning for students at risk. In H. Andrade & G. Cizek (Eds.), *Handbook of formative assessment* (pp. 212-232). New York: Routledge.
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18, 119-144.

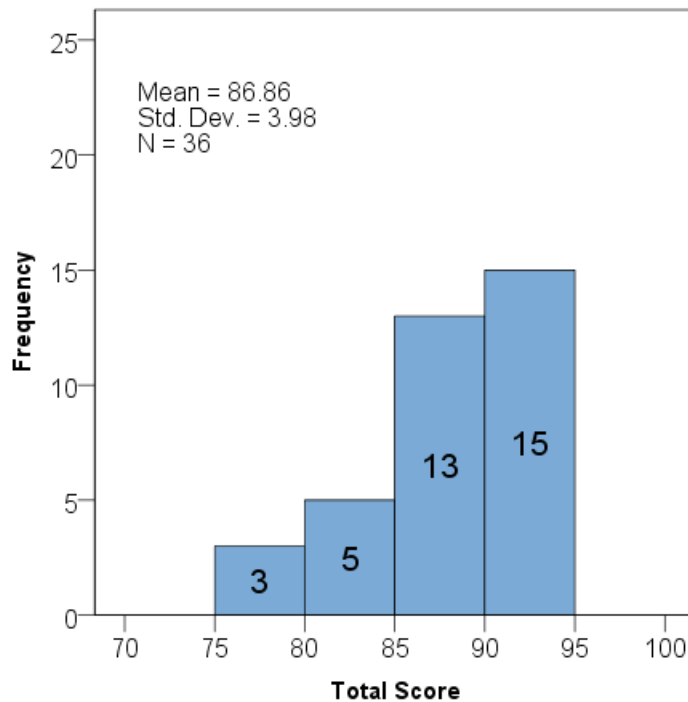


# Graph Presentations

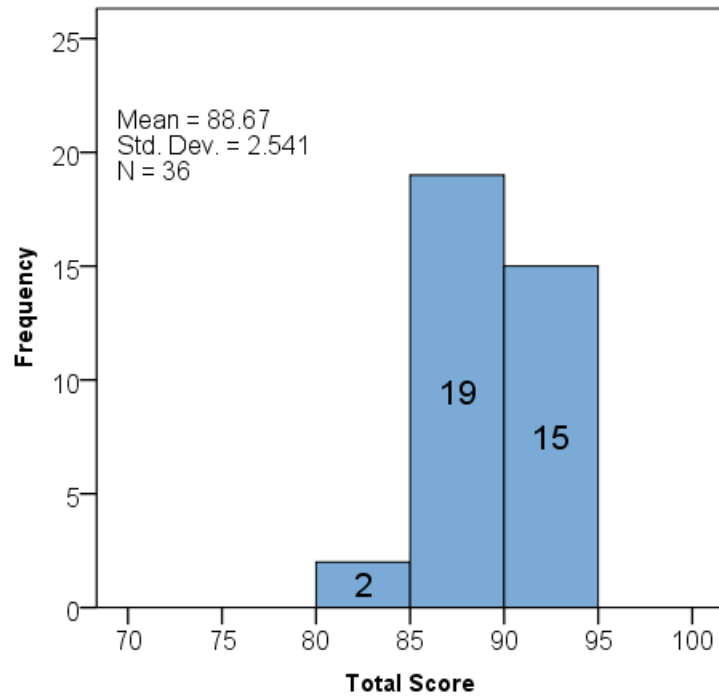
## Study 1 Graphs



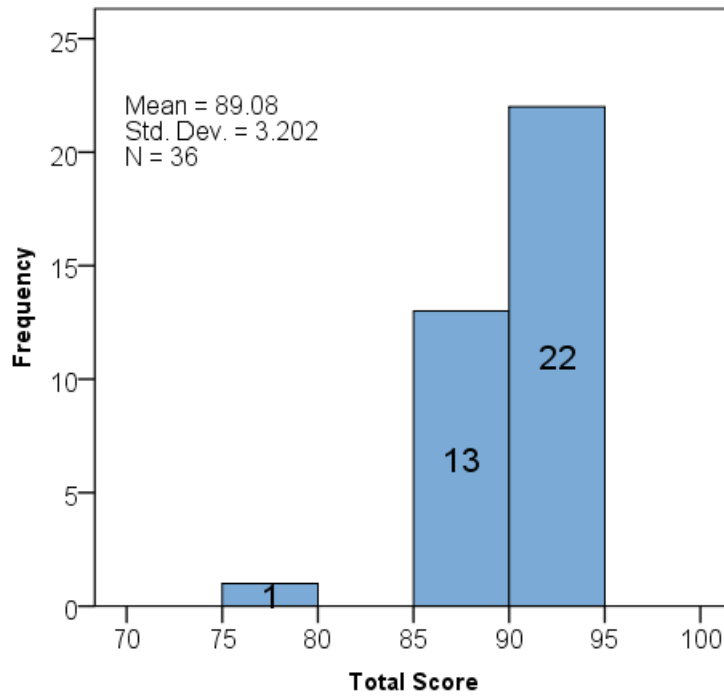
### 2012 English 003 Original Essay



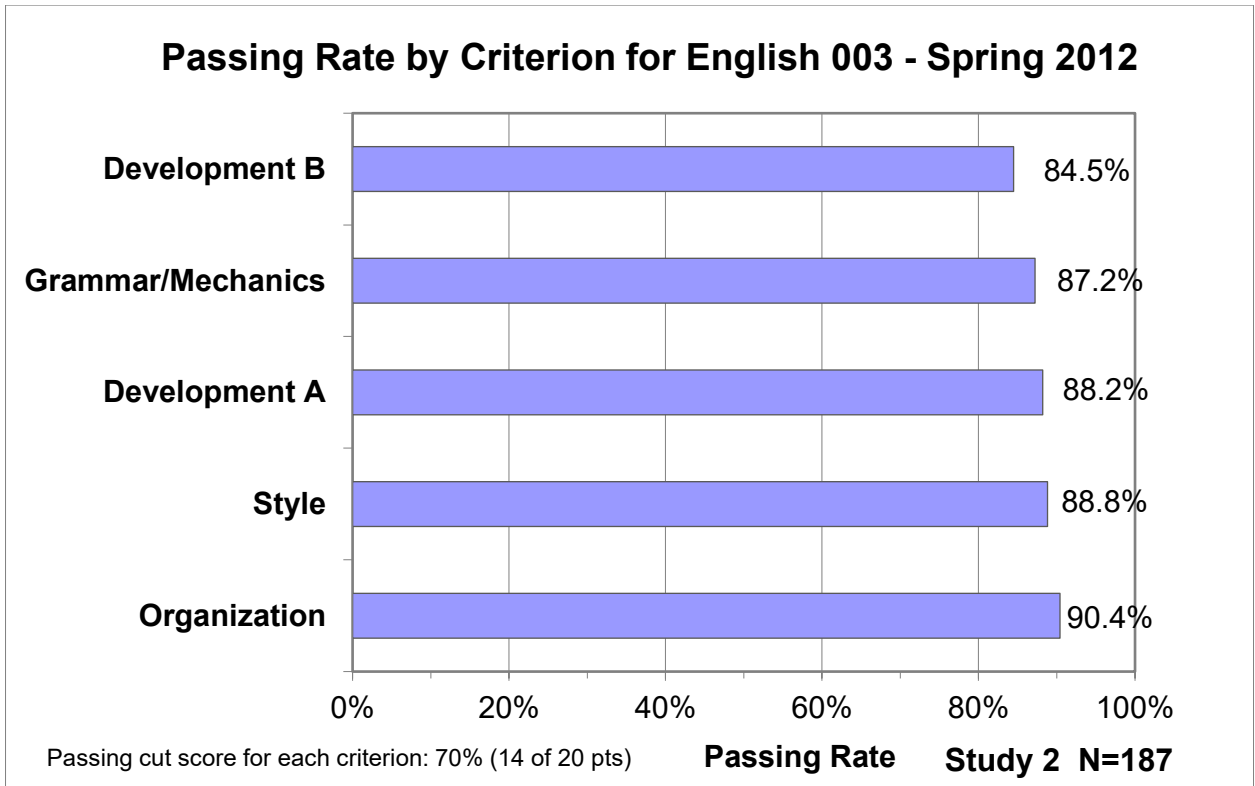
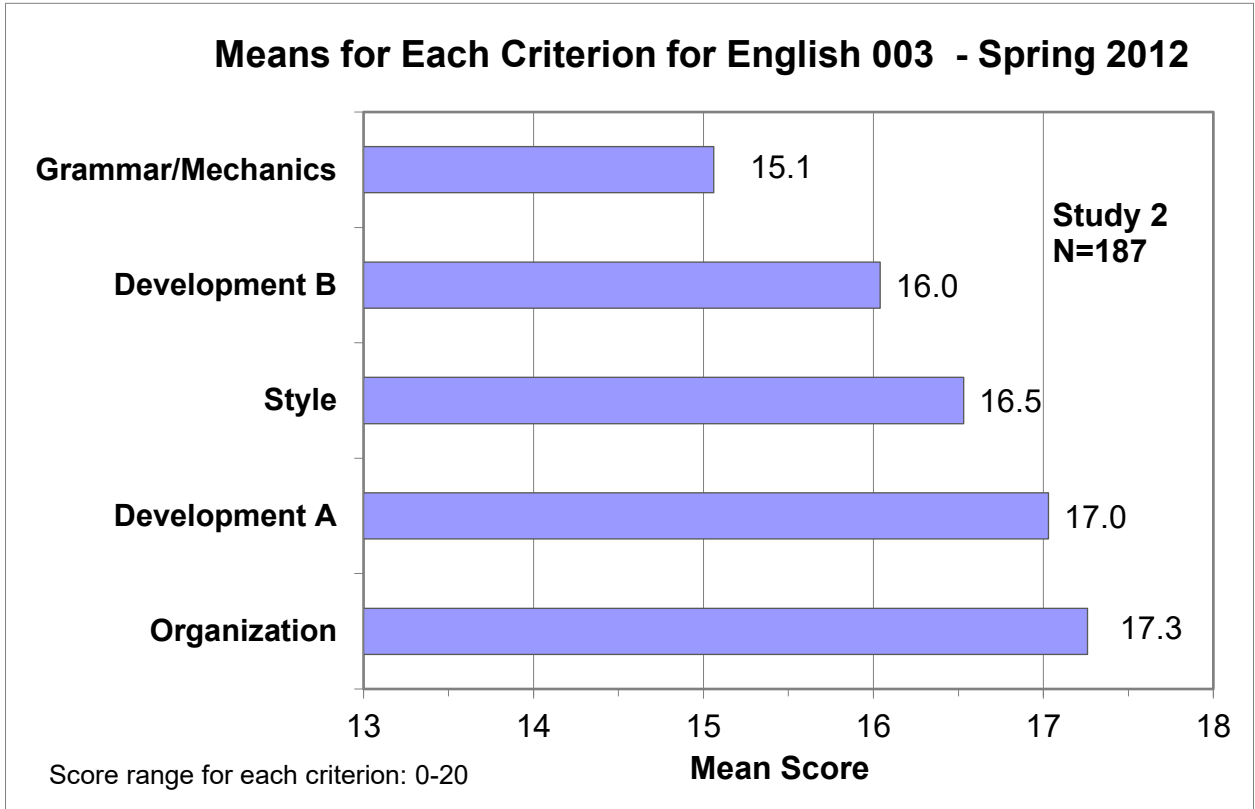
### 2012 English 003 Essay First Revision



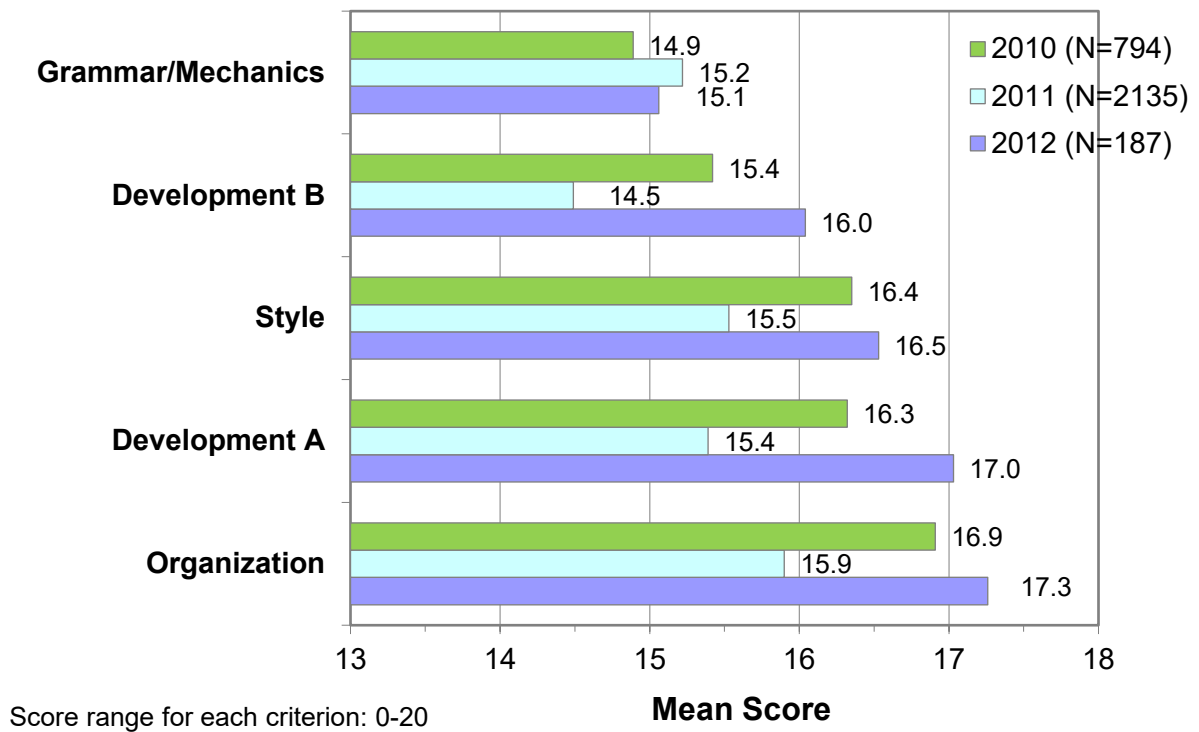
### 2012 English 003 Essay Second Revision



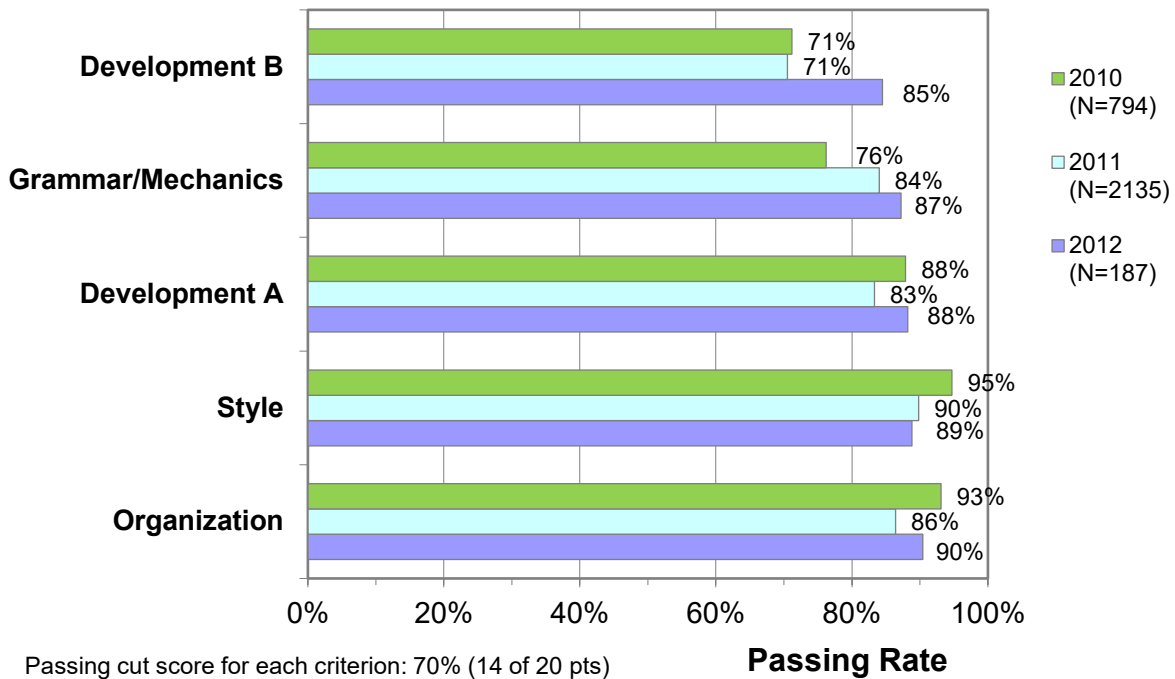
*Study 2 Graphs*



### Comparison of Criterion Means for English 003 Spring 2010, 2011, 2012



### Comparison of Passing Rate for English 003 Spring 2010, 2011 and 2012



# Appendix

ANALYTIC RUBRIC FOR THE ARGUMENTATIVE ESSAY (ENGLISH 003)  
LETTER GRADES COORDINATE WITH THIS SCHEME AS FOLLOWS: 100-90=A; 89-80=B; 79-70=C; 69-60=D; 59-BELOW F

| VALUE →<br>CRITERIA ↓   | Strong<br>18-20 pts.<br>Total 100-90 = A  | Satisfactory<br>16-17 pts.<br>89-80 = B  | Passing / Acceptable<br>14-15 pts.<br>79-70 = C  | Failing, Inadequate<br>12-13 pts.<br>69-60 = D   | Failing,<br>Seriously Flawed<br>1-11 pts.<br>59 and below  | Total Failure<br>0 pts.  | SCORE |
|---|---|--|--|--|--|--|-------|
| <b>OVERALL &amp; WITHIN PARAGRAPHS:</b><br>Weight = 20 pts  | Consistently organizes ideas logically and effectively and connects them w/ clear transitions:<br>18-20 pts.  | Usually organizes ideas logically and connects them with appropriate transitions:<br>16-17 pts.  | Organizes ideas satisfactorily but may not always connect them with transitions:<br>14-15 pts.                                 | Has limited organization of ideas:<br>12-13 pts.   | Is very disorganized; weak to no evidence of organization:<br>1-11 pts.  | Responds off topic; evidences plagiarism:<br>0 pts.  |       |
| <b>DEVELOPMENT:</b><br>A) Analysis of Argument and Supporting Evidence; Incl. Research Documentation as Relevant<br>Weight = 20 pts | Clearly identifies and insightfully analyzes important features of the argument and effectively supports the main points of the argument:<br>18-20 pts. | Clearly identifies important features of the argument and analyzes them in a generally thoughtful way and sensibly supports the main points of the argument:<br>16-17 pts. | Identifies and capably analyzes important features of the argument and supports the main points of the argument:<br>14-15 pts. | Does not identify or analyze most of the important features of the argument and offers support of little relevance or value for main points of the argument:<br>12-13 pts. | Does not understand, identify or analyze main features of the argument and/or provides little, erroneous, or illogical support for the main points of the argument:<br>1-11 pts. | Provides no evidence of ability to identify or analyze features of argument and/or provides no support; off topic; evidences plagiarism:<br>0 pts. |       |
| <b>DEVELOPMENT:</b><br>B) Opposition and Refutation Incl. Research Documentation as Relevant (as needed)<br>Weight = 20 pts.        | Clearly and fully explains opposition and completely and logically refutes it; consistently uses documentation effectively:<br>18-20 pts.               | Fully explains opposition and logically and persuasively refutes it; usually uses documentation effectively:<br>16-17 pts.   | Adequately explains opposition and gives logical refutation; adequately uses documentation:<br>14-15 pts.                      | Does not adequately explain opposition, only superficially explains. Provides some refutation; uses documentation somewhat ineffectively:<br>12-13 pts.                    | Barely explains or does not explain opposition and/or provides weak or no refutation; uses documentation ineffectively:<br>1-11 pts.   | Provides no opposition or refutation; responds off topic; does not use required documentation; evidences plagiarism:<br>0 pts.                     |       |
| <b>STYLE:</b><br>Language Control (Syntax & Diction)<br>Weight = 20 pts.  | Demonstrates superior control of language, including diction and syntactic variety:<br>18-20 pts.   | Demonstrates clear control of language, including diction and syntactic variety:<br>16-17 pts.   | Demonstrates adequate control of language, including diction and syntax but may lack syntactic variety:<br>14-15 pts.          | Uses language imprecisely and/or lacks sentence variety:<br>12-13 pts.*  | Has serious or severe and frequent problems in the use of language and sentence structure:<br>1-11 pts.  | Has no sentence structure; incorrect diction a problem; evidences plagiarism:<br>0 pts.  |       |
| <b>GRAMMAR AND MECHANICS (Punctuation):</b><br>Weight = 20 pts.   | Demonstrates superior facility with conventions (grammar, usage, and mechanics):<br>18-20 pts.  | Demonstrates facility with the conventions of standard English but may have minor flaws:<br>16-17 pts.   | Displays control of the conventions of standard written English but may have some flaws:<br>14-15 pts.                         | Has occasional major errors or frequent minor errors in grammar, usage and mechanics:<br>12-13 pts.  | Has numerous errors in grammar, usage and mechanics that interfere with meaning or result in incoherence:<br>1-11 pts.   | Has illegible and gross errors in grammar, usage and mechanics; evidences plagiarism:<br>0 pts.  |       |

Student Name/ID: \_\_\_\_\_

Total Score: \_\_\_\_\_

Rater: \_\_\_\_\_