## ASSESSMENT RESULTS FOR QCC EDUCATIONAL OUTCOME #2
### Liberal Arts Academy, Spring 2014

### Pilot Study: Liberal Arts Academy Faculty Coordinators
1. Rose-Marie Aikas- Criminal Justice
2. Renee Rhodd- Education
3. Jodie Childers- English
4. Jenny Lin Maan- Foreign Language
5. David Rothman- Academic Literacy

### GEN ED# 2 - USE ANALYTICAL REASONING TO IDENTIFY ISSUES OR PROBLEMS AND EVALUATE EVIDENCE IN ORDER TO MAKE INFORMED DECISIONS

### QCC Example Outcomes:
- a. Distinguish the problem or question from a proposed solution or answer
- b. Differentiate between facts, assumptions, and conclusions in the formulation of a proposed solution or answer
- c. Evaluate the quality of evidence
- d. Describe and compare the way questions, issues, or problems are formulated within various fields of study

### GEN ED OUTCOME ADDRESSED IN ASSESSMENT:
(See Rubric: Attached)
- a. Distinguish the problem or question from a proposed solution or answer
- b. Take a specific position on an issue, taking into account the complexities of the issue at hand (thesis/argument)
- c. Illustrate an understanding of the issue/problem addressed using relevant content.
- d. Support claims with evidence taken from multiple sources

### EVIDENCE/MEASURABLE DATA
- **Describe artifacts reviewed**: 50 student essays were collected from across the academic disciplines that make up the Liberal Arts Academy. 5 essays were collected from 10 LA courses from the following disciplines: English/Academic Literacy, History/Political Science, Criminal Justice, Foreign language, Education and Health Sciences.
- **No. of artifacts**: 5 essays
- **No. of sections**: 5 essays

<table>
<thead>
<tr>
<th>Course Code</th>
<th>Course Title</th>
<th>No. of Essays</th>
</tr>
</thead>
<tbody>
<tr>
<td>BE 112</td>
<td>Composition Workshop-</td>
<td>5</td>
</tr>
<tr>
<td>HI 128</td>
<td>Growth of American Civ II</td>
<td>5</td>
</tr>
<tr>
<td>IS 151</td>
<td>Health of the Nation</td>
<td>5</td>
</tr>
<tr>
<td>En 101</td>
<td>English Composition 1</td>
<td>5</td>
</tr>
<tr>
<td>EN 412</td>
<td>American Literature II</td>
<td>5</td>
</tr>
<tr>
<td>LS-223</td>
<td>Workshop in Reading and Writing</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>for Spanish Heritage Speakers</td>
<td></td>
</tr>
<tr>
<td>CRIM 101</td>
<td>Criminal Justice 1</td>
<td>5</td>
</tr>
<tr>
<td>CRIM 102</td>
<td>Criminal Justice 2</td>
<td>5</td>
</tr>
<tr>
<td>EDU 101</td>
<td>Education 101</td>
<td>5</td>
</tr>
<tr>
<td>PLSCI 101</td>
<td>Political Science 101</td>
<td>5</td>
</tr>
</tbody>
</table>
## RESULTS

The collected data from the rubric assessment was categorized in a number of different ways (see attached Excel spreadsheet- ‘Liberal Arts Assessment Project Data’).

Note: the Excel spreadsheet ‘The Liberal Arts Assessment Key’ arranges the fifty essays by both course and first and second readers.

1. **Comparing Reader Reliability and scores by domain**

   This data is organized into five spreadsheet pages from Essays 1-10, 11-20, 21-30, 31-40 and 41-50. There were only four discrepant scores out of the fifty essays evaluated. The rater reliability was 92%. The four discrepant scores are evident on the spreadsheet as 3rd reads are highlighted in yellow. In the case of a discrepancy in scoring, only the third reader’s score counted. In our analysis, the rare case of discrepant scoring took place when readers were evaluating fact-driven essays, which did not explicitly take a position or cite their sources of evidence. Some readers gave more credit for implicit arguments and sourcing.

2. **Overall Average Score and Average Score by Domain**

   The average individual reader score for an essay was 10.583333333. Once again, the highest possible score across the four domains would be 16 (4/4/4/4) and the lowest possible score would be 4 (1/1/1/1). Thus, the result places the average score between the two (minimally acceptable) score and the three (competent) score. The average score is closer to a competent rating of three, as a rating hitting a competent average would be at 12.0 compared with a rating hitting a minimally acceptable average would be 8.0.

   If we examine the average score by domain, we can see that the student essays received the highest average score in the first domain, ‘Issue/Problem’, with the average score being 2.82291. The second highest average score came in the third domain of ‘Development’ where the average score was 2.70833. The third highest average score fell in the second domain of ‘Student Perspective’, with the average score being 2.64583. Finally, the lowest average score was found in the fourth domain, ‘Evidence’, with an average score of 2.58333.

   It is far from surprising that the highest average score came in the first domain of Issue/Problem as identifying the central issue in a given essay assignment may not be as challenging for some students as offering a clear thesis or providing relevant support with clear evidence. The fact that the lowest average score came in the fourth domain of ‘Evidence’ may follow from the fact that some of the given assignments did not specify that students must cite multiple sources (see discussion of caveats below).

3. **Comparing Upper-Level Course Essays with Lower-Level Course Essays**

   Student essays from upper-level courses, EN 412: American Literature II and LS-223: Workshop in Reading and Writing for Spanish Heritage Speakers, had a higher average score than those from lower-level courses. The average score for upper-level essays was 11.59375 compared with the overall average score of 10.58333. Upper-level
essays had higher average scores across all four domains. In the ‘Issue/Problem’ domain, while the average score was 2.8229, the average score for upper-level essays was 3.1875. In the Student’s Perspective’ domain, while the average score was 2.6458, the average score for upper-level essays was 2.78125. With the third domain, ‘Development’, while the average score was 2.7083, the average score for the upper-level course essays was 2.9375. Finally, in the fourth domain of ‘Evidence’ the average score was 2.5833 while the average score for upper-level course essays was 2.6875.

4. Comparing Results by Academic Discipline

Given that our readers are mostly limited to evaluating student work within their own given academic discipline, we thought it might be interesting to compare how student essays were scored by discipline. The Political Science essays outscored all of the other essays by discipline, even though the one Political Science course essays were collected from was a 100-level course. The average score for the Political Science essays was 13.666, well above the competent essay baseline of 12.0. After political science came the essays taken from Foreign Language, with an average score of 12.2, just above the ‘competent’ mark. This may partially be the result of the fact that the papers collected from Foreign Language were from an upper-level course. The same may be said of the essays collected from English, where half of the essays came from an upper-level course. The average score for the essays from English was 12.0. The Education, Health and Criminal Justice average results were quite similar. The Education average score was 9.916. The Health essays average was 9.7692 and the Criminal Justice essays had an average score of 9.0. Finally, not surprisingly, the essays collected from developmental writers in Academic literacy had the lowest average score, at 7.8.

<table>
<thead>
<tr>
<th>STATUS – degree to which students have met Gen. Ed. Outcome</th>
<th>Superior (16 points)</th>
<th>Competent (12 points)</th>
<th>Minimally Acceptable (8 points)</th>
<th>Poor (4 points)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>20 essays (40%)</td>
<td>21 essays (42%)</td>
<td>9 essays (18%)</td>
<td></td>
</tr>
</tbody>
</table>

RATIONALE (Briefly explain rating above)

We adapted the AAC&U Critical Thinking VALUE Rubric and the AAC&U Written Communication Rubric, combining and modifying both so that the final rubric aligned more precisely with our specific learning outcome. Our rubric is divided into four domains (Issue-Problem/Student’s Perspective/Development/Evidence. Our scoring scale or each domain ranged from a high of 4 points to a low of 1 point. Thus, a perfect score from one reader would be 16 points (4/4/4/4) and the lowest score possible from each individual reader would be 4 points (1/1/1/1).

(see Rubric-Attachment A)

4 score; Represents an essay that fully meets the criteria of a superior essay
Qualifying descriptors in rubric articulations: comprehensive/compelling

3 score: Represents a competent essay
Qualifying descriptors in rubric articulations: coherent/ample/relevant
| DISCUSSION OF THIS PROCESS: What did you learn and what would you recommend? | 2 score: Represents a minimally acceptable essay  
Qualifying descriptors in rubric articulations: simplistic/some understanding  
1 score represents a poor essay  
Qualifying descriptors in rubric articulations: unclear/little or no understanding  
- While a few essays received one score of 16 from a given reader, none of the essays reached the ‘Superior’ level with both readers giving a score of 16.  
- 20 essays, or 40% of the essays, received a two-reader average score above 12 points.  
- 21 essays, or 42% of the essays received combined scores between the baseline for minimally acceptable (8 points) and 12 points, where the next level begins. Finally, 9 essays, or 18% of those evaluated, received combined scores below 8 points, which placed them in the Poor range. |
|---|---|
| **1. Process Timeline** | November 2013: Meeting with Jane Hindman, Ian Beckford and LAA faculty coordinators to review the goals and implementation of the project  
- Gen Ed Outcome #2: Analytical reasoning was chosen as the outcome we would focus on. A discussion of the GEN ED objectives led us to the conclusion that analytical reasoning is a particularly relevant skill across the many academic disciplines represented in the Liberal Arts Academy.  
December 2013: Collection of 50 essays from across the academic disciplines represented in Liberal Arts Academy  
- Each faculty coordinator solicited sets of student essays from LAA faculty. Sets of five essays were collected from faculty teaching the following courses (Fall 2013 classes):  
  BE 112: Composition Workshop  
  HI 128: Growth of American Civilization II  
  IS 151: Health of the Nation  
  En 101: English Composition 1  
  EN 412: American Literature II  
  LS-223: Workshop in Reading and Writing for Spanish Heritage Speakers  
  CRIM 101: Criminal Justice 1  
  CRIM 102: Criminal Justice 2  
  EDU 101: Education 101  
  PLSCI 101: Political Science 101  
The essays were de-identified and classified by number rather than by name.  
January 2014: The LAA faculty coordinators created a rubric based on the desired outcomes of GEN ED #2: Use analytical reasoning to identify issues or problems and evaluate evidence in order to make informed decisions.  
February 2014- Five faculty members from across the Liberal Arts were asked to take part in the scoring of our fifty collected essays.  
- Jilani Warsi - Academic Literacy  
- Melissa Dennehy - English  
- Arancha Borrachero - Foreign Language  
- Jennifer Maloy - Academic Literacy |
March 2014- Two norming sessions were held with LAA faculty coordinators working with our rubric to score sample essays. Ian Beckford served as an advisor at both sessions.

- Faculty coordinators discussed and carefully reviewed the articulations laid out in the rubric. Scoring discrepancies were discussed and the coordinators gained a better understanding of the scoring scale.
- The 50 essays were then distributed with each reader given a packet which contained the rubric; a set of 12-13 essays to score; a clearly laid out scoring sheet.
- The plan was for each essay to be scored by two readers. Any essay with a discrepant score (defined as containing any non-adjacent scores) would then be given to a third reader.

April 2014 - The 50 essays were scored over a two week period
- LAA faculty coordinators met to organize the scored essays.

May 2014 - Liberal Arts Faculty Coordinators reviewed the data and wrote up their findings.

Caveats:

a. The sampling of collected essays were somewhat homogeneous in terms of course level. They were mostly limited to 100 level courses and remedial-level classes. The exceptions were five essays taken from English 412: American Literature II, and five taken from LS-223: Workshop in Reading and Writing for Spanish Heritage Speakers. In a future study, it would make sense to collect a wider essay sampling from across course levels.

b. There was wide variation in the extent to which essay assignments prompted students to evaluate multiple sources of evidence and to cite their sources. These criteria must be made explicit in the essay assignments. The faculty coordinators should only collect essays based on assignments, which demand analytical reasoning and the evaluation of multiple sources. Put simply, students can only be assessed on what they have been prompted to do.