Differences in Student Writing Ability as a Function of Student Characteristics at One Texas 4-Year University

Jeff Roberts
Director of Assessment
Sam Houston State University

Presented at the 2016 Southwest Educational Research Association Annual Meeting
Introduction

- Written communication remains a key learning objective for today’s college students

- Employers report a strong desire for institutions to emphasis written communication

- However, only 27% of employers believe that recent graduates are well-prepared with regard to written communication
Some researchers also hold a negative perception of student writing ability

- The Spellings Commission noted that students were graduating without necessary skills in written communication

- These findings were echoed within the (in)famous book *Academically Adrift*
  - Arnum & Roska, 2011
Review of the Literature

- Historical/meta-analysis of literature regarding writing assessment
  - Anson, 2010, Anson & Lyles, 2011; Behizadeh & Englehard, 2011

- Studies of writing assessment theory and practice
  - Anson, 2006; Gallagher, 2010

- Studies in which student writing ability was examined
  - Alan & Driscoll, 2014; Good et al., 2012
The first step to address critics and improve student writing is to assess student writing accurately.

Written communication is of particular interest to Texas institutions

- The Texas Higher Education Coordinating Board includes student written communication as a core learning objective
  - THECB, 2015
Purpose of the Study

This study originated out of one university's efforts to assess student writing
- Nardone et al., 2013; Roberts et al., 2014.

The goals of the original writing assessment were two-fold:
- Evaluate the effectiveness of writing-enhanced courses
- Collect base-line data regarding student-writing ability
  - Nardone et al., 2013; Roberts et al., 2014.

This study builds upon that work, utilizing archived data to answer additional questions regarding student writing

Furthermore, it serves as a potential model for other writing assessments
Significance of the Study

- Results from this study will be used to help improve student writing ability within the studied institution.

- The results of this study may be used by other researchers to assess and improve student writing at their institutions.

- The results from this study may be helpful to other Texas institutions in their assessment of student writing as part of their State-mandated core curriculum assessment efforts.
Research Questions

- What is the difference in the student performance on an end-of-experience student writing assessment as a function of student race?

- What is the difference in the student performance on an end-of-experience student writing assessment as a function of student gender?
Delimitations

- Sample encompassed only Junior- and Senior-level students enrolled within 4000-level Writing Enhanced courses at one university in south-east Texas.

- Data were only gathered from the Spring 2013 academic semester.
  - Therefore only represent a snapshot of student writing ability.
Limitations

The nature of the sample pool means that the results may not be generalizable to different student populations, different institutions, and different locations.

As these data were only gathered from one academic semester, any relationships or differences identified may represent anomalies, and not be reflective of actual trends over time.
Assumptions

- It is assumed that any errors within the dataset are random and not specific to any one group or variable
  - Data were previously collected and verified by the author of the study; therefore, minimal errors are anticipated
  
  - Authentic student writing artifacts were used and are assumed to represent the best possible examples of student work
  
  - The rubric was developed by interdisciplinary group of faculty with expertise in student writing, and is therefore assumed to have content-related validity
    - Banta & Palomba, 2015; Bridges et al., 2013
Method
Nonexperimental, causal comparative research design

- Design allows for the use of existing data
- Does not allow for the manipulation of the examined variables
Participants

- Junior- and senior-level students enrolled in 4000-level writing enhanced courses during the Spring 2013 semester

- A stratified random sampling process was used to select student artifacts for analysis
  - 395 student artifacts were used for scoring

- Sample was representative of the size and diversity of the university’s student population
Locally developed writing rubric with four domains:
- Ideas/Critical Thinking/Synthesis
- Style
- Organization
- Conventions

Each artifact received a separate score for each of the four domains using a 4-point scale

Two raters evaluated each artifact independently

Third rater introduced when scores were out of agreement
Score Reliability

ICC’s were calculated to determine the level of inter-rater agreement.

<table>
<thead>
<tr>
<th>Category Area</th>
<th>Intraclass Correlation for Average Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ideas/Critical Thinking/Synthesis</td>
<td>.69 – Good</td>
</tr>
<tr>
<td>Style</td>
<td>.65 – Good</td>
</tr>
<tr>
<td>Organization</td>
<td>.64 – Good</td>
</tr>
<tr>
<td>Conventions</td>
<td>.58 – Fair</td>
</tr>
<tr>
<td>Overall Total</td>
<td>.80 – Excellent</td>
</tr>
<tr>
<td>Overall Average</td>
<td>.80 – Excellent</td>
</tr>
</tbody>
</table>

According to Cicchetti (1994), ICC agreement:
> .40 = poor agreement
.40-.59 = fair agreement
.60-.74 = good agreement
.75 <= excellent agreement
Student writing scores that were derived from locally-developed writing rubric unique to that institution were used in this study.

Therefore, no attempt is made to generalize the findings of this study beyond its circumstances.
Results

- Parametric Multivariate Analysis of Variance (MANOVA) procedures were used to determine whether any differences could be observed in student writing scores as a function of student race and student ethnicity.
  - The use of a parametric procedure was justified as the majority of assumptions were met (Field, 2009) for both research questions.
    - Skewness and kurtosis coefficients were within the ranges of normality (+/-3; Onwuegbuzie & Daniel, 2001).
    - Box’s Test’s of Equality of Covariance were violated.
    - Levene’s Test of Equality of Error Variances were met.
No statistically significant differences were observed in student writing as a function of either student race or of student gender.

- Student Writing as a function of race
  - Wilks’ Λ = .97, p = .56

- Student Writing as a function of gender
  - Wilks’ Λ = .99, p = .65
## Descriptive Statistics for Student Overall Writing Scores by Student Demographic Characteristic

<table>
<thead>
<tr>
<th>Student Demographic Characteristic</th>
<th>n</th>
<th>M %</th>
<th>SD %</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Race</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>259</td>
<td>2.69</td>
<td>0.64</td>
</tr>
<tr>
<td>Black</td>
<td>51</td>
<td>2.53</td>
<td>0.65</td>
</tr>
<tr>
<td>Hispanic</td>
<td>56</td>
<td>2.54</td>
<td>0.56</td>
</tr>
<tr>
<td>Other</td>
<td>28</td>
<td>2.48</td>
<td>0.53</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>143</td>
<td>2.59</td>
<td>0.65</td>
</tr>
<tr>
<td>Female</td>
<td>251</td>
<td>2.66</td>
<td>0.61</td>
</tr>
</tbody>
</table>
If an institution is doing an adequate job of preparing its students, one would expect to see equitable results regardless of student type.

Therefore, the lack of statistically significant differences in writing scores as a function of race or gender may be interpreted as a positive result.

To confirm these results, more study is needed:
- Eliminate competing theories
- Eliminate the possibility of error
Discussion

- Flaws could exist with the rubric or assessment methodology.

- Findings could be the result of an inadequate sample size, thus limiting the observed differences between the groups.

- Weaker students, regardless of race or gender, may not persist to the junior and senior years, thus limiting the observed differences between the groups.
Discussion

- Sample size could be increased.

- Rubric could be validated through cross-type and cross-institutional scoring.
  - Use the Rubric to score artifacts from entry-level students
  - Use the rubric to score artifacts from another institution
Finally, equity does not mean quality...

More information is needed to determine whether the level of student performance observed within this study was sufficient for end-of-experience students.
No magic bullet exists for assessing student learning. No one test, measure, or rubric will ever provide all the answers needed by faculty, staff, and administrators to improve student learning.

Additionally, improvements do not occur over night, but take time and intentionality.
To improve student writing institutions must ultimately have assessments that provide reliable and valid data that are meaningful to them.

This study represents one such attempt by a Texas 4-year university.

It is hoped that it will inspire others to assess student writing at their own institutions.
Questions?

Jeff Roberts
Director of Assessment
Sam Houston State University

jeff.roberts@shsu.edu
936–294–1859