Panel: The VALUE of Quality Degrees
AAC&U Annual Meeting
January 2014

Mo Noonan Bischof
Assistant Vice Provost
mabischof@wisc.edu

Jocelyn Milner
Director of Academic Planning and
Institutional Research
jlmilner@wisc.edu
Helping students *live* the Wisconsin Idea through active acquisition & application of knowledge

**Wisconsin Experience**

“Wisconsin Experience” captures:
- Substantial research experiences that generate knowledge and analytical skills
- Global and cultural competencies and engagement
- Leadership and activism opportunities
- Application of knowledge in the “real world”

**Percent of Graduates with One or more Wisconsin Experience Activity**

<table>
<thead>
<tr>
<th>Activity</th>
<th>2012-13 Graduates</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Graduates</td>
<td>89%</td>
</tr>
<tr>
<td>Targeted Minority Graduates</td>
<td>93%</td>
</tr>
</tbody>
</table>

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%
Our Learning Community

21,615 employees…
2,177 faculty
1,635 instructional academic staff
1,261 research academic staff
5,291 graduate assistants

42,820 students …
29,118 undergraduates
9,183 graduate students
2,774 professional students
1,745 Non-degree students

Annually:
7,400 new undergraduates
29,500 enrolled undergraduates
6,500 Bachelor’s degree graduates
13 academic schools/colleges
distributed responsibility and governance
~500 academic programs, all levels
134 Bachelor’s level degree programs

Helping students live the Wisconsin Idea through active acquisition & application of knowledge
Institutional-level learning goals, assessments

Program-level learning goals, assessments

Program-level learning goals, assessments

Program-level learning goals, assessments

Program-level learning goals, assessments

Incorporates WI-X and ELO's learning goals and assessments
AAC&U VALUE
Rubric Project

- 25 faculty
- cross-disciplinary representation
- focus on faculty engagement

**Scorers**

**Rubrics**

**Artifacts**

AAC&U VALUE written communication rubric

- "Value-added" approach to compare first year students (FYR) and students near graduation (NGR)
- Goal was to collect 350 artifacts at each level
- Identified 52 courses that had high numbers of FYR and NGR and seemed likely to have a suitable writing assignment
- 22 courses (41 instructors) had a suitable assignment and agreed
- Invited 2450 students to submit artifacts
- Collected 451 submissions.

**Written Communication VALUE Rubric**

**Definition**

Written communication in the development and expression of ideas in writing. Written communication involves learning to work in many genres and styles. It can involve working with many different writing strategies, and meeting many criteria and standards. Written communication involves dealing through literate experience across different domains.

**Steps**

1. **Capture**
2. **Rubric**
3. **Artifact**
4. **Assessment**

**Dimensions**

<table>
<thead>
<tr>
<th>Dimensions</th>
<th>Capture</th>
<th>Rubric</th>
<th>Artifact</th>
<th>Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Context and Purpose for Writing</td>
<td>- Does the task accurately represent the assignment?</td>
<td>- Does the task accurately represent the assignment?</td>
<td>- Does the task accurately represent the assignment?</td>
<td>- Does the task accurately represent the assignment?</td>
</tr>
<tr>
<td>Clear, Concise, and Complete</td>
<td>- Does the writing clearly and concisely communicate the intended message?</td>
<td>- Does the writing clearly and concisely communicate the intended message?</td>
<td>- Does the writing clearly and concisely communicate the intended message?</td>
<td>- Does the writing clearly and concisely communicate the intended message?</td>
</tr>
<tr>
<td>Content Development</td>
<td>- Does the writing effectively develop the intended message?</td>
<td>- Does the writing effectively develop the intended message?</td>
<td>- Does the writing effectively develop the intended message?</td>
<td>- Does the writing effectively develop the intended message?</td>
</tr>
<tr>
<td>Genre and Disciplinary Connection</td>
<td>- Does the writing appropriately connect to the discipline?</td>
<td>- Does the writing appropriately connect to the discipline?</td>
<td>- Does the writing appropriately connect to the discipline?</td>
<td>- Does the writing appropriately connect to the discipline?</td>
</tr>
<tr>
<td>Coherence and Clarity</td>
<td>- Does the writing show a clear, logical development of ideas?</td>
<td>- Does the writing show a clear, logical development of ideas?</td>
<td>- Does the writing show a clear, logical development of ideas?</td>
<td>- Does the writing show a clear, logical development of ideas?</td>
</tr>
<tr>
<td>Control of Voice and Mechanics</td>
<td>- Does the writing show a clear, logical development of ideas?</td>
<td>- Does the writing show a clear, logical development of ideas?</td>
<td>- Does the writing show a clear, logical development of ideas?</td>
<td>- Does the writing show a clear, logical development of ideas?</td>
</tr>
</tbody>
</table>
ABOUT SCORERS

• 1.5 day workshop in June 2013
• Guests: Terrel Rhodes and Ashley Finley
• Set ground rules
• 3 structured rounds intended to get faculty familiar with the rubric and to “test” scorer agreement
• Asked faculty to think beyond their field/discipline
• Two scorers per rubric on 4 point scale with “4” representing appropriate level for graduating senior and “1” representing an appropriate level for an incoming student
• Each scorer rated about 40 artifacts
• Discussion revealed challenge with the 4-point scale and what is “mastery”
• Good news: High level of engagement with faculty and the quality of the discussions was important as we move forward with discussions around the use of rubrics for assessment at the program level.

The Ground Rules

- This is not grading.
- Think globally about student work and about the learning skill. Think beyond specific disciplinary lenses or content.
- We are not changing the rubric (today).
- Our work is time sensitive. Go with your instinct.
- Start with 4 and work backwards.
- Pick one performance benchmark per criterion. Avoid “.5”.
- Zero does exist. Assign “0” if work does not meet benchmark “1” performance level. N/A exists. Assign “not applicable” if the student work is not intended to meet a particular criterion.

Credit: Terrel Rhodes and Ashley Finley, AAC&U
Exercise

- With nearby colleagues, review the results. What do they tell us? What conclusions do you draw?

- If these were your findings, what would you do next?

- Do you use rubrics at the program level or campus-wide? What challenges have you dealt with? What successes have you had?

Table 1. Overall Results for All Artifact Scores

<table>
<thead>
<tr>
<th>Rubric Dimension</th>
<th>Student Group</th>
<th># of Artifacts</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Zmw Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Context</td>
<td>Nearly Graduating</td>
<td>213</td>
<td>2.95</td>
<td>0.95</td>
<td>3.05*</td>
</tr>
<tr>
<td></td>
<td>First Year</td>
<td>237</td>
<td>2.77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Content</td>
<td>Nearly Graduating</td>
<td>213</td>
<td>2.79</td>
<td>0.96</td>
<td>4.68*</td>
</tr>
<tr>
<td></td>
<td>First Year</td>
<td>237</td>
<td>2.48</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genre</td>
<td>Nearly Graduating</td>
<td>211</td>
<td>2.69</td>
<td>0.88</td>
<td>2.65*</td>
</tr>
<tr>
<td></td>
<td>First Year</td>
<td>235</td>
<td>2.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sources</td>
<td>Nearly Graduating</td>
<td>190</td>
<td>2.61</td>
<td>0.99</td>
<td>1.54</td>
</tr>
<tr>
<td></td>
<td>First Year</td>
<td>225</td>
<td>2.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Syntax</td>
<td>Nearly Graduating</td>
<td>213</td>
<td>2.82</td>
<td>0.84</td>
<td>2.16*</td>
</tr>
<tr>
<td></td>
<td>First Year</td>
<td>237</td>
<td>2.69</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Zmw score is from the Mann Whitney U-Test. Zmw scores >1.96 indicate that the two groups are significantly different at p=0.05.
Table 1. Distribution of Combined Scores - Written Communication Rubric

<table>
<thead>
<tr>
<th>Percent of Scores</th>
<th>First-Year Students</th>
<th>Nearly Graduating Students</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.8</td>
<td>2.6</td>
</tr>
<tr>
<td>2</td>
<td>27.3</td>
<td>22.6</td>
</tr>
<tr>
<td>3</td>
<td>51.0</td>
<td>44.4</td>
</tr>
<tr>
<td>4</td>
<td>30.3</td>
<td>17.5</td>
</tr>
</tbody>
</table>

Table 2. Inter-Scorer Reliability (Krippendorff’s Alpha Co-efficient)

<table>
<thead>
<tr>
<th>Agreement Strength</th>
<th># of scorer pairs</th>
<th>% of pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strong: + 0.7 to 1.0</td>
<td>8</td>
<td>5%</td>
</tr>
<tr>
<td>Fair : + 0.3-0.69</td>
<td>47</td>
<td>28%</td>
</tr>
<tr>
<td>Weak: + 0.0-0.29</td>
<td>43</td>
<td>25%</td>
</tr>
<tr>
<td>Negative : &lt;0.0</td>
<td>70</td>
<td>42%</td>
</tr>
</tbody>
</table>

The range for Krippendorff’s alpha co-efficient is +1 to -1 with +1 signaling perfect agreement and scores less than zero signally systematic disagreement beyond that expected by chance.

Krippendorff’s alpha coefficients were calculated for all 168 unique pairs of scorers. Values >0.7 are considered to represent “strong” agreement; 0.3 to 0.7 to represent fair agreement, 0 to 0.3 to represent weak agreement. Overall, 67% of scorer pairs showed weak agreement or systematic disagreement.
What did we learn?

Rubric approaches in our environment are 

Expensive → TIME, MONEY 
Worth it → FACULTY ENGAGEMENT 
Worth it → PROGRAM LEVEL 

Contact Information 
Mo Noonan Bischof, Assistant Vice Provost, University of Wisconsin-Madison, mabischof@wisc.edu 
Jocelyn Milner, Director of Academic Planning and Institutional Research, University of Wisconsin-Madison, jmilner@wisc.edu 
More about our project: http://apir.wisc.edu/valuerubricproject.htm