# Table of Contents

- Executive Summary .......................................................... 3
- Overview ............................................................................ 3
- Artifact Collection ............................................................ 5
- AAC&U VALUE Rubric Design ............................................. 5
- Scoring Session ................................................................. 7
- Results .............................................................................. 7
- Key Findings ..................................................................... 12
- Recommendations ............................................................. 13
- Discussion of Results and Future Plans ......................... 13
Executive Summary

At the University of Maine (UMaine), the general education curriculum makes up one third of a student’s academic experience. The purpose of general education assessment is to evaluate how well our students are meeting the general education learning outcomes and to better understand areas needing improvement. There are five overarching general education areas and nine discrete general education areas:

1. Western cultural tradition
   a. Social context and institutions
   b. Cultural diversity and international perspectives
   c. Population and the environment
   d. Artistic and creative expression

2. Ethics

3. Quantitative literacy

4. Writing

5. Science foundations

In January 2020 and January 2021, the Cultural Diversity and International Perspectives block was assessed. Faculty used modified AAC&U VALUE rubrics to score digital copies of student artifacts and submit the scores in Google Forms. A group of 27 participants scored 187 artifacts. The information from this report will be used to inform the general education curriculum and future assessment activities.

Key Findings:

- Inter-rater reliability percentages indicate a low to moderate agreement among raters.
- Results indicate that for three of the rubric categories, nearly a third or more of the artifacts did not show evidence of meeting a benchmark level.
- From the artifacts with evidence of meeting the rubric categories, the majority of scores were at a level 1 or level 2.
- There were no significant differences between scores of first-years/sophomores versus juniors/seniors.

Recommendations:

- More rubric training is needed.
- Consider the definition of the Cultural Diversity and International Perspectives general education category.
- Greater effort needed to align outcomes with assessments.

Overview

Introduction

The general education curriculum intends to help develop “broadly educated persons who can appreciate the achievements of civilization, understand the tensions within it, and contribute to resolving them” (UMaine Catalog, 2019). There are many ways for a student to fulfill the requirements of the general education curriculum, and it makes up one third of a student’s academic experience. Five broad categories comprise the general education curriculum:

1. Human Values and Social Contexts

General Education Assessment Report

a. Western Cultural Tradition
b. Social Context and Institutions
c. Cultural Diversity and International Perspectives
d. Population and the Environment
e. Artistic and Creative Expression

2. Ethics
3. Quantitative Literacy
4. Writing
5. Science

Additionally, a capstone experience specific to the student’s major is required. More details can be found at https://umaine.edu/facultysenate/committees/general-education-committee/.

In January 2020 and January 2021, the Cultural Diversity and International Perspectives block of the general education curriculum was assessed. As stated in UMaine’s General Education documentation, “A course included in the Cultural Diversity or International Perspectives category satisfies one or more of the following criteria: (a) it places primary emphasis on the experiences, perspectives, and cultural work of one or more groups who are not dominant within a particular culture; (b) it has a primary goal encouraging students to become aware of the diversity of American culture and to discover their roles within that diversity; or (c) it places primary emphasis on the relationships among or within different cultures in the past or present; (d) it introduces students to a culture other than their own through an intermediate or advanced course in the language of that culture.

Students completing the Cultural Diversity or International Perspectives general education area will be able to do at least one of the following:

1. Recognize the experiences, perspectives, and cultural values of one or more groups who live within a culture different than their own.
2. Describe the diversity of American culture and reflect on their personal roles within that diversity.
3. Identify and assess how different cultures have related to each other either in the past or the present.
4. Achieve intermediate or advanced mastery of a language other than English.

Organization and Reporting

The general education assessment scoring session was organized by the General Educational Committee of the Faculty Senate and the Office of Institutional Research and Assessment. The rubrics used to assess student artifacts were adapted from the AAC&U VALUE rubrics. Faculty were organized into teams based on the general education classes they taught. They designed rubrics aligned to the nine general education areas using criteria from the VALUE rubrics. While only two general education areas have been assessed so far, the rubrics to assess each of the areas have been developed and will be used in future scoring sessions.

A statistically representative sample of artifacts from lower division courses and upper division courses was randomly selected by the Office of Institutional Research and Assessment to be evaluated by faculty reviewers. Individual faculty members whose courses were randomly selected were contacted via email to submit student artifacts that aligned to the general education area being assessed. Ten students were then randomly selected (by the Office of

---

2 The original scoring session was scheduled for January 2020. Due to inclement weather the session was quickly moved online and the participation dropped. Due to COVID, a subsequent session was not scheduled until January 2021, in which the remainder of the artifacts were scored.
Institutional Research and Assessment) from each group of artifact submissions. Any identifying information of the student or the course was redacted by the Office of Institutional Research and Assessment.

After completing the scoring sessions, the results were collected and analyzed by the Office of Institutional Research and Assessment and reported back to the General Education Committee of the Faculty Senate.

Artifact Collection

Student artifacts were collected from 19 course sections from a stratified random sample from the Registrar’s list of course sections. Representative of lower division and upper division proportions of general education courses, 14 lower division courses (74%) and 5 upper division courses (26%) were randomly sampled. Ten students were randomly selected from each course; however, due to small class size and tardy submissions, less than ten artifacts were collected from some of the courses. In total, 187 artifacts were collected for the scoring session. There were 137 artifacts from lower division courses and 50 artifacts from upper division courses.

The artifacts were embedded course assignments collected from courses that met the general education designation as determined by the Undergraduate Program Curriculum Committee (UPCC). All artifacts were redacted so the student and course would be unidentifiable.

Rubric Design

The rubric used for the scoring event was developed by a team of faculty using the AAC&U VALUE rubrics. For the Cultural Diversity and International Perspectives general education block, five rubric categories were assessed. The categories (represented in the rubric below) are as follows:

1. Civic Identity and Commitment
2. Evidence selecting and using information to investigate a point of view or conclusion
3. Student’s Position (perspective, thesis/hypothesis)
4. Application of Ethical Perspectives/Concepts
5. Global Self-Awareness
6. Personal and Social Responsibility

### Cultural Diversity or International Perspectives

#### Preamble

A course included in the Cultural Diversity or International Perspectives category satisfies one or more of the following criteria: (a) it places primary emphasis on the experiences, perspectives, and cultural work of one or more groups who are not dominant within a particular culture; (b) it has a primary goal encouraging students to become aware of the diversity of American culture and to discover their roles within that diversity; or (c) it places primary emphasis on the relationships among or within different cultures in the past or present; (d) it introduces students to a culture other than their own through an intermediate or advanced course in the language of that culture.

#### Student Learning Outcomes

Students completing the Cultural Diversity or International Perspectives general education area of will be able to do at least one of the following: Gen Ed Committee Reviewed 2/17/12

1. Recognize the experiences, perspectives, and cultural values of one or more groups who live within a culture different than their own.
2. Describe the diversity of American culture and reflect on their personal roles within that diversity.
3. Identify and assess how different cultures have related to each other either in the past or the present.
4. Achieve intermediate or advanced mastery of a language other than English.

<table>
<thead>
<tr>
<th>Description</th>
<th>Capstone</th>
<th>Level 3</th>
<th>Level 2</th>
<th>Benchmark</th>
</tr>
</thead>
<tbody>
<tr>
<td>Civic Identity and Commitment</td>
<td>Provides evidence of experience in civic-engagement activities and describes what she/he has learned about her or himself as it relates to a reinforced and clarified sense of civic identity and continued commitment to public action.</td>
<td>Provides evidence of experience in civic-engagement activities and describes what she/he has learned about her or himself as it relates to a growing sense of civic identity and commitment.</td>
<td>Evidence suggests involvement in civic-engagement activities is generated from expectations or course requirements rather than from a sense of civic identity.</td>
<td>Provides little evidence of her/his experience in civic-engagement activities and does not connect experiences to civic identity.</td>
</tr>
<tr>
<td>Evidence Selecting and using information to investigate a point of view or conclusion</td>
<td>Information is taken from source(s) with enough interpretation/evaluation to develop a comprehensive analysis or synthesis.</td>
<td>Information is taken from source(s) with enough interpretation/evaluation to develop a coherent analysis or synthesis.</td>
<td>Information is taken from source(s) with some interpretation/evaluation, but not enough to develop a coherent analysis or synthesis.</td>
<td>Information is taken from source(s) without any interpretation/evaluation.</td>
</tr>
<tr>
<td>Student's position (perspective, thesis/hypothesis)</td>
<td>Specific position (perspective, thesis/hypothesis) is imaginative, taking into account the complexities of an issue. Limits of position (perspective, thesis/hypothesis) are acknowledged. Others' points of view are synthesized within position (perspective, thesis/hypothesis).</td>
<td>Specific position (perspective, thesis/hypothesis) takes into account the complexities of an issue. Others' points of view are acknowledged within position (perspective, thesis/hypothesis).</td>
<td>Specific position (perspective, thesis/hypothesis) acknowledges different sides of an issue.</td>
<td>Specific position (perspective, thesis/hypothesis) is stated, but is simplistic and obvious.</td>
</tr>
<tr>
<td>Application of Ethical Perspectives/Concepts</td>
<td>Student can independently apply ethical perspectives/concepts to an ethical question, accurately, and is able to consider full implications of the application.</td>
<td>Student can independently apply ethical perspectives/concepts to an ethical question, accurately, but does not consider the specific implications of the application.</td>
<td>Student can apply ethical perspectives/concepts to an ethical question, independently (to a new example) and the application is inaccurate.</td>
<td>Student can apply ethical perspectives/concepts to an ethical question with support (using examples, in a class, in a group, or a fixed-choice setting) but is unable to apply ethical perspectives/concepts independently (to a new example.).</td>
</tr>
</tbody>
</table>
## Scoring Session

This scoring session was unique from past scoring sessions. It was originally due to take place in January 2020. However, due to inclement weather, the session was quickly moved online and participation dropped significantly. Not enough artifacts were scored to create a report. A subsequent session was planned for spring 2020. However, due to COVID-19, the session did not take place until January 2021. The remainder of the artifacts were scored at this time.

In total, twenty-seven UMaine faculty and staff signed up to participate in the scoring session in January 2021. Participants were asked to independently view a training video and score one calibration artifact prior to the event. Scorers were then directed to watch a short video explaining the scoring before beginning the process. Each participant had an individual folder with 10-11 artifacts to score. All scores were collected via Google Forms.

The rubric used during the scoring event consisted of four levels: 4 (capstone), 3 (milestone), 2 (milestone), and 1 (benchmark). If an artifact did not meet a level 1 on a rubric category or showed no evidence related to that criteria, the scorers were instructed to indicate a score of “0.” For example, a level 1 in the rubric category of Global Self-Awareness calls for evidence that the student “identifies some connections between an individual’s personal decision-making and certain local and global issues.” If the paper clearly contained elements of wrestling with global and local issues, but the student did not make connections to personal decision making, the artifact would be given a score of “0” for that rubric category.

## Results

Our general education assessment scoring process is modeled after AAC&U’s VALUE Institute’s rubrics and methodology. While rubrics more commonly ascend upward in level, the VALUE rubrics descend from level 4 to level 1. The purpose of this is to emphasize an “assets-based---versus deficit-focused---approach to assessment of student learning.” Scorers are trained to assume students are capable of attaining the highest level and working down from there.

After the scoring event was completed, the data were downloaded from the Google Forms spreadsheet and cleaned for analysis. The results were analyzed by rubric category and student level. Although the rubric categories are numeric (4,3,2,1), the data are categorical and do not reflect a true scale. For example, the distance between a level 1 and a level 2 may not be equal to the distance between a level 3 and a level 4. The data are therefore presented as percentages of scores per rubric category. To determine a final score for each rubric category among artifacts that had two scorers,
two general rules were applied (as adopted from AAC&U’s method). If an artifact was given a score of 0 by at least one scorer, the 0 was retained and the other score was deleted. Secondly, if an artifact was given a score of two different whole numbers (4,3,2,1), the score was averaged and rounded up to the nearest whole number. For example, if an artifact was given scores of 3 and 4 on a rubric category, the final score for that category would be a 4.

**Inter-rater Reliability**

Several measures of inter-rater reliability were calculated, a process modeled after that used by AAC&U to prepare their yearly VALUE Reports (Drezek McConnell & Rhodes, 2017). Each artifact was scored by two raters. For the purposes of determining inter-rater reliability all ratings, including 0s, were included. As stated above, 0s meant that there wasn’t enough evidence of either alignment or student ability. In the table below, exact agreement reports the percentage of artifacts scored exactly the same by the two raters while adjacent agreement reports the percentage of agreement within one rubric level of each other (for example if one rater scored 1 while the other scored 2). From these results we see that exact agreement was generally low: approximately 36% averaged across rubric categories. However, the relationship was notably stronger when we looked at the percentage in relatively close agreement: approximately 70% averaged across rubric categories. The remainder of the statistics in the table adjust for chance agreement in different ways. Values closer to 1 are better in each case. The weighted Cohen’s Kappa takes into account the degree of disagreement between raters. The weighted Brennan-Prediger Kappa coefficient controls for the number of scoring levels (0-4) and the Gwet’s AC coefficient considers how uniform the rating distribution was for each rubric category. Categories with uniform distributions are harder to rate (less consensus) than ones with more peaked distributions (more consensus). While the Cohen’s Kappa results indicate relatively low agreement, the Brennan-Predifer and Gwet’AC results indicate low to moderate levels of agreement. In general, absolute agreement was rare, but raters tended to be in the same ballpark with their application of the rubric criteria.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Exact Agreement (%)</td>
<td>49.01%</td>
<td>33.11%</td>
<td>32.45%</td>
<td>36.42%</td>
<td>27.81%</td>
<td>36.42%</td>
</tr>
<tr>
<td>Adjacent Agreement (%)</td>
<td>72.85%</td>
<td>73.51%</td>
<td>74.83%</td>
<td>65.56%</td>
<td>63.58%</td>
<td>70.20%</td>
</tr>
<tr>
<td>Weighted Cohen’s Kappa</td>
<td>0.18</td>
<td>0.25</td>
<td>0.27</td>
<td>0.22</td>
<td>0.15</td>
<td>0.18</td>
</tr>
<tr>
<td>Weighted Brennan-Prediger</td>
<td>0.44</td>
<td>0.45</td>
<td>0.49</td>
<td>0.23</td>
<td>0.30</td>
<td>0.37</td>
</tr>
<tr>
<td>Weighted Gwet’s AC</td>
<td>0.73</td>
<td>0.52</td>
<td>0.57</td>
<td>0.52</td>
<td>0.44</td>
<td>0.60</td>
</tr>
</tbody>
</table>

**Scoring Results**
The results of the scoring session were analyzed in three different ways: a.) evidence of skills emphasized in rubric vs. no evidence, b.) score distributions per rubric category, and c.) score distributions per student level.

The Cultural Diversity and International Perspectives rubric contains five categories:

1. Civic Identity and Commitment
2. Evidence selecting and using information to investigate a point of view or conclusion
3. Student’s Position (perspective, thesis/hypothesis
4. Application of Ethical Perspectives/Concepts
5. Global Self-Awareness
6. Personal and Social Responsibility

Scorers determined to what extent an artifact met the criteria in each rubric category and gave it a corresponding score (4,3,2,1,0). Many artifacts were given a score of “0”, indicating that there was no evidence of reaching at least a benchmark level (1) or there was no evidence of alignment between the artifact and the rubric category. The chart below depicts the percentage of artifacts that displayed evidence of the skills or knowledge required by at least a benchmark level in each rubric category (bottom percentage-blue) versus artifacts that did not show evidence of achieving benchmark or alignment (top percentage-orange).

**Evidence vs. No evidence**

The stacked bar charts below depict the score distribution across rubric categories when all course types and student levels are considered together. Percentages of students scoring at each level (4,3,2,1) per rubric category are shown descending from a level 4 to a level 1. Represented in the chart are the percentages of artifacts showing evidence of achieving at least a benchmark level-- scores of 0 were not included in this analysis. The table below the chart gives a more nuanced description of the counts and percentages for each rubric category based on the total of artifacts with
evidence versus those with no evidence. From the artifacts with evidence of meeting the rubric categories, the majority of scores were at a level 1 or 2.

### Score Distributions per Rubric Category

[Chart showing score distributions per rubric category]

#### Score Distributions per Student Level

The following chart and table give a breakdown of scores per rubric category by class standing (freshman & sophomore, junior & senior). A more detailed table of the breakdown of percentages per category follows this chart.
Score Distributions per Student Level in each Rubric Category

Rubric Categories

<table>
<thead>
<tr>
<th>Fr./So.</th>
<th>Jr./Sr.</th>
<th>Fr./So.</th>
<th>Jr./Sr.</th>
<th>Fr./So.</th>
<th>Jr./Sr.</th>
<th>Fr./So.</th>
<th>Jr./Sr.</th>
<th>Fr./So.</th>
<th>Jr./Sr.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Civic Identity</td>
<td>Evidence Selecting and Using Information</td>
<td>Student Position</td>
<td>Application of Ethical Perspectives</td>
<td>Global Self-awareness</td>
<td>Personal and Social Responsibility</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Capstone

<table>
<thead>
<tr>
<th>First-Years and Sophomores</th>
<th>Count</th>
<th>%</th>
<th>Milestones</th>
<th>Count</th>
<th>%</th>
<th>Benchmark</th>
<th>Count</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Civic Identity</td>
<td>1</td>
<td>2</td>
<td>7</td>
<td>17</td>
<td>14</td>
<td>33</td>
<td>20</td>
<td>48</td>
</tr>
<tr>
<td>2. Evidence Selecting and Using Information</td>
<td>5</td>
<td>8</td>
<td>19</td>
<td>29</td>
<td>30</td>
<td>46</td>
<td>11</td>
<td>17</td>
</tr>
<tr>
<td>3. Student's Position</td>
<td>2</td>
<td>3</td>
<td>19</td>
<td>29</td>
<td>23</td>
<td>35</td>
<td>21</td>
<td>32</td>
</tr>
<tr>
<td>4. Application of Ethical Perspectives</td>
<td>6</td>
<td>12</td>
<td>16</td>
<td>31</td>
<td>14</td>
<td>27</td>
<td>15</td>
<td>29</td>
</tr>
<tr>
<td>5. Global Self-awareness</td>
<td>4</td>
<td>7</td>
<td>12</td>
<td>20</td>
<td>20</td>
<td>34</td>
<td>23</td>
<td>39</td>
</tr>
<tr>
<td>6. Personal and Social Responsibility</td>
<td>1</td>
<td>2</td>
<td>9</td>
<td>16</td>
<td>20</td>
<td>36</td>
<td>25</td>
<td>45</td>
</tr>
</tbody>
</table>

Juniors and Seniors

<table>
<thead>
<tr>
<th>First-Years and Sophomores</th>
<th>Count</th>
<th>%</th>
<th>Milestones</th>
<th>Count</th>
<th>%</th>
<th>Benchmark</th>
<th>Count</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Civic Identity</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>8</td>
<td>20</td>
<td>41</td>
<td>24</td>
<td>49</td>
</tr>
<tr>
<td>2. Evidence Selecting and Using Information</td>
<td>18</td>
<td>21</td>
<td>33</td>
<td>38</td>
<td>25</td>
<td>29</td>
<td>11</td>
<td>13</td>
</tr>
<tr>
<td>3. Student's Position</td>
<td>13</td>
<td>14</td>
<td>20</td>
<td>22</td>
<td>32</td>
<td>35</td>
<td>27</td>
<td>29</td>
</tr>
<tr>
<td>4. Application of Ethical Perspectives</td>
<td>7</td>
<td>11</td>
<td>10</td>
<td>16</td>
<td>23</td>
<td>37</td>
<td>23</td>
<td>37</td>
</tr>
<tr>
<td>5. Global Self-awareness</td>
<td>5</td>
<td>7</td>
<td>13</td>
<td>17</td>
<td>23</td>
<td>30</td>
<td>35</td>
<td>46</td>
</tr>
<tr>
<td>6. Personal and Social Responsibility</td>
<td>3</td>
<td>5</td>
<td>10</td>
<td>17</td>
<td>21</td>
<td>36</td>
<td>25</td>
<td>42</td>
</tr>
</tbody>
</table>
Key Findings

1. **Inter-rater reliability percentages indicate a low to moderate agreement among raters.**

   Although calibration sessions took place before the official scoring began, inter-rater reliability remained low to moderate. Consequently, we cannot say there was strong agreement on how artifacts scored against the rubric. Given the low reliability and feedback from the scorers regarding the confusing language in components of the rubric, we conclude that the Cultural Diversity and International Perspectives rubric should be evaluated by the Academic Affairs and General Education Committee of the Faculty Senate. Further training on the rubric, whether it is revised or stays in its current form, is necessary.

2. **Results indicate that for three of the rubric categories, nearly a third or more of the artifacts did now show evidence of meeting a benchmark level.**

   The “Evidence of Each Rubric Category” chart shows many of the artifacts seemed to have no relation to this set of general education learning outcomes or corresponding rubric. Without decent alignment of the artifacts to the rubric, the scoring session can seem more random than helpful for understanding the general education curriculum.

3. **From the artifacts with evidence of meeting the rubric categories, the majority of scores were at a level 1 or 2.**

   When removing all scores of “0” from the analysis, the majority of artifacts were scored at a level 2. Results from the Western Cultural Tradition 2018 and Social Contexts and Institutions 2019 scoring sessions yielded similar findings.

4. **There were not significant differences between scores of first-years/sophomores versus juniors/seniors.**

   One might expect the scores of student artifacts from the junior/senior level to score higher than the scores of student artifacts from the first-years/sophomores. However, there are no strong patterns indicating this is the case.

Recommendations

1. **More rubric training is needed.**

   Similar to findings from other general education scoring sessions, the Cultural Diversity and International Perspectives rubric clearly posed challenges during the scoring events. Discussion during the January 2021 session focused on intended meaning of rubric categories and semantics. More attention should be paid to preparing scorers and reaching a shared agreement on the meaning of the rubric categories and descriptions. The large percentage of artifacts containing no evidence of addressing the outcome suggests more effort is needed to encourage faculty to align their courses and assessments to this, and other, General Education outcomes.

2. **Consider the definition of the Cultural Diversity and International Perspectives general education category.**

   In a followup session after the scoring was complete, several faculty expressed being confused between what is expected in a Cultural Diversity and International Perspectives course. There was some skepticism about whether or
not the rubric captures what the learning outcomes intend to measure. One small but actionable suggestion was to change the category “Global self-awareness” to simply “Global awareness.”

**Discussion of Results and Future Plans**

Assessment data collected from the general education assessment session will be discussed within the Academic Affairs and General Education Committee of the Faculty Senate and then shared internally and publicly as appropriate to encourage discussion about programmatic changes.

The results from this year are in line with findings from past scoring sessions in other general education categories. Questions about alignment and confusion around the learning outcomes and rubrics are ongoing themes of these sessions. It is clear from this report and from the faculty conversation following the scoring session that there is a need for greater awareness about the general education learning outcomes and corresponding rubrics. Plans are underway to disseminate the rubrics more widely, in hopes that they will eventually be used in assignment design or else altered to be more usable.

Assessment of the general education curriculum is ongoing, however, a Faculty Senate motion was passed in May 2021 to pause the general education assessment process just long enough to consider how to strengthen our current process or propose an alternative one. This is because the findings from three scoring sessions point to the need to make this process more meaningful for the improvement of our general education curriculum. Specifically, considering how to facilitate better alignment of the general education outcomes with the corresponding courses will be a significant part of the discussion. In Fall 2021, the General Education Subcommittee of the Academic Affairs Faculty Senate committee will begin discussions on the current rubrics, the current assessment process, and possible alternative methods of assessing our general education curriculum. Additionally, a larger general education committee is being formed and will soon be commissioned by Provost Volin to examine our general education curriculum at large. Undoubtedly, the conversations of both committees will inform each committee’s work over the next year. Dependent on decisions made in Fall 2021, the next general education category to be assessed in May 2022 is Population & Environment.