General Education Assessment Review

Other World Civilizations

Hirosuke Honda
Institutional Research Analyst

Center for Planning and Institutional Research
9/30/2010
Introduction
In 1998, the State University of New York’s Board of Trustees passed a resolution that was the basis for forming the Provost’s Advisory Task Force on the assessment of student learning outcomes. This group established the framework for the SUNY Assessment Initiative, which mandated assessment of all academic majors and general education programs across the State University system, with the focus on student learning outcomes. As a result, task forces were convened at colleges across the state with the responsibility of developing an assessment plan detailing procedures and guidelines as well as a timeline for implementation. A group consisting of SUNY faculty, students, and representatives from System Administration began reviewing campus assessment plans during the 2001-02 academic year, with the first year of campus-based general education assessment beginning the following year.

Empire State College’s original GEAR assessment plan outlined a schedule in which one assessment would be conducted in 2003-04, three in 2004-05, and eight assessments in 2005-06. That plan was later revised and changed to reflect current practices, which requires the college to conduct four General Education Assessment Reviews per year, on a three-year cycle.

Rating Instrument
Prior to the inaugural assessment in each general education area, Empire State College convened task forces consisting of faculty teaching courses in these respective disciplines. These faculty members were asked to design rubrics based on SUNY learning outcomes that could be used to rate samples of student work against specified objectives. These rubrics consist of explicitly stated criteria describing student work that “does not meet,” “approaches,” “meets,” or “exceeds” student learning objectives. In instances where raters perceived that a student work sample is not relevant to one of the rubric’s objectives, a fifth category of “not applicable,” or NA, is utilized. For reporting purposes each score is coded numerically from 1-4, where 1=“does not meet,” 2=“approaches,” 3=“meets,” and 4=“exceeds.” These rubrics were approved by the Undergraduate Student Policies Committee as well as administrators within the Office of Academic Affairs prior to their use.

Prior to the inaugural GEAR in Other World Civilizations at Empire State College, faculty members designed the Other World Civilizations rubric. This rubric consisted of two student learning objectives:

1. Demonstrate knowledge of the development of the Features of non-Western Civilization of the history, institutions, economy, society, culture, etc., of Western civilization.

2. Relate the development of Western civilization to that of Knowledge of World History Concepts of the world.

Procedure
Prior to each assessment, staff members in the Center for Planning and Institutional Effectiveness (C-PIE) collected student work samples in designated general education areas from the Center for Distance Learning’s online archives. Upon collection of materials, a team of three raters convene for the assessment to review and rate student work samples against established scoring rubric(s). Prior to the assessment itself, a norming session lead by a C-PIE staff was conducted with faculty. The purpose of this norming session is to familiarize faculty with the scoring rubric and to build a group consensus. This entails distributing a sample of student work and having all faculty members rate its content against the rubric’s multiple objectives. Participating faculty were then asked to share their ratings. In instances
where discrepancies between raters existed, faculty discussed their ratings with the goal of building consensus or at least narrowing differences among raters.

**Student Work Sample**
The initial GEAR in Other World Civilizations was conducted in May of 2007. A total of 32 work samples from students studying at various centers across the state and 71 samples from the Center for Distance Learning’s online archives were collected for the assessment (total 103). The second GEAR in Other World Civilizations was conducted in May 2010. A total of 100 work samples were collected from CDL archives for this assessment. Since each student work sample was rated by two faculty raters, there were a total of 206 ratings in 2007, and 200 ratings and one sample rated once in 2010 (total 201).

**Reliability**
There are two types of reliability calculated for General Education Assessment Reviews. One is internal consistency, which measures the extent to which objectives on a rubric are rated similarly. If a rubric has two objectives and one is rated as “does not meet” and the other is rated as “exceeds,” and this is consistent across student work samples, the rubric would lack internal consistency. The reliability statistic used to measure internal consistency is Cronbach’s alpha. An alpha level of 0.70 or higher is considered to be acceptable. The Cronbach’s alpha level was 0.90 for the 2007 GEAR in Other World Civilizations and 0.70 for the 2010 assessment. Both of these levels were well above the threshold for acceptability indicating that the Empire State College Other World Civilizations GEAR rubric possesses good internal consistency.

The other type of reliability calculated for these sets of data was interrater reliability. Interrater reliability is the measure of the extent to which multiple judges agree on their ratings. If raters generally agree and assign a similar rating to samples of student work for individual expectations, the rubric would be considered reliable. Interrater reliability is estimated using Intraclass correlations (ICCs). Again, an alpha level of 0.70 or higher is considered to be acceptable. In the 2007 assessment, agreement between raters was above the threshold of acceptability on the “Features of non-Western Civilization” objective, but below the threshold on the “Knowledge of World History Concepts” objective. In 2010, the ICC on both of the rubric’s learning objectives was above the 0.70 threshold, with intraclass correlations ranging from 0.93 to 0.95, indicating that the Other World Civilizations rubric was being utilized by faculty raters in a similar manner for the 2010 GEAR in Other World Civilizations. Results are depicted below in Table 1.

<table>
<thead>
<tr>
<th>Learning Objectives</th>
<th>2007</th>
<th>2010</th>
</tr>
</thead>
<tbody>
<tr>
<td>Features of non-Western Civilization</td>
<td>0.856</td>
<td>0.953</td>
</tr>
<tr>
<td>Knowledge of World History</td>
<td>0.899</td>
<td>0.926</td>
</tr>
</tbody>
</table>
**Results**

In the 2010 GEAR in Other World Civilizations, the percentage of faculty ratings of “meets” or “exceeds” for student work was 52.3% for the “Features of non-Western Civilization” learning objective and 38.4% for the “Knowledge of World History” objective. The mean score for the former objective was 2.49, and the latter was 2.37. These numbers indicate that on average student work samples fell in the “approach” category the “Features of non-Western Civilization” learning objective and the “meet” category for the “Knowledge of World History Concepts” objective.

In comparison with the 2007 assessment, the percentage of ratings of “meets” or “exceeds” for both learning objectives increased: “Features of non-Western Civilization” from 35.4% to 52.3%, and “Knowledge of World History” from 36.3% to 38.4%. A Chi square analysis was performed on this data to determine whether the differences between the assessments in Other World Civilizations were statistically significant. The results revealed that significant differences existed on both objectives. Complete results for both assessments appear below in Table 2.

**Analysis and Discussion**

The rubric used in both the 2007 and 2010 General Education Assessment Reviews in Other World Civilizations has two distinct student learning objectives: 1) Demonstrate knowledge of the development of the distinctive features of the history, institutions, economy, society, culture, etc., of at least one non-Western civilization; and 2) Demonstrate knowledge of a broad outline of world history. These objectives were written by faculty prior to the 2007 assessment and were determined to be in alignment with SUNY learning outcomes in the area of Other World Civilizations.

Results from the 2010 GEAR in Other World Civilizations revealed that student samples were overall approaching standards on both of the rubric’s objectives, as evident by mean scores of 2.49 on the “Features of non-Western Civilization” learning objective and 2.37 on the “Knowledge of World History Concepts” objective. These numbers indicate that the student work samples meet the first learning objective, and approach the second objective.

Comparisons between the 2007 and 2010 General Education Assessment Reviews in Other World Civilizations exhibited that the percentage of rating of “meets” or “exceeds” for “Features of non-Western Civilization” increased from 35.4% to 52.3%. In contrast, the percentage of the two categories remained consistent in “Knowledge of World History.”

---

**Table 2: GEAR Other World Civilizations Comparison Between 2007 and 2010 Assessments**

<table>
<thead>
<tr>
<th>Learning Objective</th>
<th>Valid N</th>
<th>Does not meet</th>
<th>Approaches</th>
<th>Meets</th>
<th>Exceeds</th>
<th>NA</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Features of non-Western Civilization 2007**</td>
<td>161</td>
<td>41 25.5%</td>
<td>63 39.1%</td>
<td>14 8.7%</td>
<td>12</td>
<td>2.19</td>
<td>0.92</td>
<td></td>
</tr>
<tr>
<td>Features of non-Western Civilization 2010*</td>
<td>172</td>
<td>17 9.9%</td>
<td>65 37.8%</td>
<td>78 45.3%</td>
<td>12</td>
<td>2.49</td>
<td>0.77</td>
<td></td>
</tr>
<tr>
<td>Knowledge of World History 2007***</td>
<td>44</td>
<td>10 22.7%</td>
<td>18 40.9%</td>
<td>10 22.7%</td>
<td>6 13.6%</td>
<td>129</td>
<td>2.27</td>
<td>0.97</td>
</tr>
<tr>
<td>Knowledge of World History 2010**</td>
<td>78</td>
<td>5 6.4%</td>
<td>43 55.1%</td>
<td>26 33.3%</td>
<td>4 5.1%</td>
<td>123</td>
<td>2.37</td>
<td>0.68</td>
</tr>
</tbody>
</table>

Note: Percentages are based on the number of valid cases. Mean ratings based on the following scale: 1=“does not meet,” 2=“approaches,” 3=“meets,” 4=“exceeds.” Ratings of “NA” were excluded from mean calculations.

*χ² (3, N=333) = 19.90, p < .001
**χ² (3, N=122) = 10.79, p < .05
In the 2007 GEAR in Other World Civilizations, approximately 30.8% of work samples were collected from faculty located at regional centers across the state, while the remaining samples were chosen from the Center for Distance Learning’s online archives. All work samples from the 2010 GEAR in Other World Civilizations were collected from CDL archives. This slight difference in sampling must be kept in mind when making direct comparisons between the two assessments.

During the assessment session, raters began to notice that there was a variation amongst the guidance of writing assignments as well. Some assignments provided clear guidance regarding the assigned paper formatting, outlines, citations, and sometimes criteria of evaluation. Contrarily other assignments provided a short paragraph of what students were supposed to write. The raters found that students tended to perform well under clear guidance of an assignment. This observation may provide a great insight for mentors’ instructional development training and workshops.

Considering the manageability of assessment, the C-PIE staff drew student writing samples that are from five to ten pages so that raters can process an expected total number of the samples in the two-day session. However, there are many studies that assigned shorter than several pages or longer than ten page papers. Some other studies required multiple mini papers and discussion postings in an electronic board as well. Recognizing the variety of assignments within the academic subject, mentors were encouraged to discuss how they design assignments that should better reflect the learning objectives in the rubric.

One approach that could strengthen the assessment process in the general education area of Arts at Empire State College would be to implement an electronic portfolio system that could serve as a repository for student work. All students at the college could designate work samples from each study to be displayed in this portfolio. This would allow work samples to be collected from undergraduate students enrolled at centers across the state, rather than primarily from students enrolled at the Center for Distance Learning. This sample would be much more reflective of the undergraduate population and would better serve the college in terms of generalizability of assessment findings. An electronic portfolio also would allow staff from C-PIE to select student work samples that are more relevant to each learning objective or at least most of the rubric’s learning objectives.
Appendix A: Empire State College Other World Civilizations Rubric

Please fill in the circle that best reflects your assessment of each of the following expectations for student’s work in Other World Civilizations. Refer to the rubric for the appropriate evaluative criteria for each student learning objective.

<table>
<thead>
<tr>
<th>Does not meet</th>
<th>Approaches</th>
<th>Meets</th>
<th>Exceeds</th>
<th>N/A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demonstrate knowledge of the development of the distinctive features of the history, institutions, economy, society, culture, etc., of at least one non-Western civilization OR</td>
<td>Does not identify features of the history, institutions, economy, society, and/or culture of the non-Western Civilization. Does not distinguish among intellectual currents and major concepts and is not able to identify periods of cultural development, distinctive style and/or great works.</td>
<td>The work describes some features of the history, institutions, economy, society, and/or culture of the non-Western Civilization. Minimally distinguishes among intellectual currents and major concepts and is able to identify periods of cultural development, distinctive style and/or great works.</td>
<td>The work describes the major features of the history, institutions, economy, society, and/or culture of non-Western Civilization and shows some insight as to the interactions between these distinctive features. Distinguishes among intellectual currents and major concepts and is able to identify and describe periods of cultural development, distinctive style and/or great works.</td>
<td>Demonstrates comprehension of the major features of the history, institutions, economy, society, and/or culture of non-Western Civilization and shows insight into the interactions between these distinctive features. Clearly distinguishes among intellectual currents and major concepts and is able to identify, describe and explain periods of cultural development, distinctive style and/or great works.</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>Demonstrate knowledge of a broad outline of world history</td>
<td>Not able to provide any analysis of periods of historical significance or identify periods of change. Does not make connections across time periods.</td>
<td>Able to provide some analysis of periods of historical significance as well as identify periods of change, but without a broad world context. Makes limited connections across time.</td>
<td>Able to provide an adequate analysis of periods of historical significance as well as describe periods of change within the context of world history. Makes basic connections across time periods.</td>
<td>Able to provide a thorough, clear and critical analysis of periods of historical significance as well as describe and explain periods of change within the context of world history. Makes sophisticated connections across time periods.</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

Comments: