

**New Mexico State University – Las Cruces
Assessment of General Education Student Learning Outcomes
2015-16 Final Report**

**Submitted by the
Committee for the Assessment of Student Learning in General Education
(CASL-GE)**

Ryan Goss – Plant & Environmental
Sciences, Chair
Greg Armfield – Communication Studies
Susan Beck – Library
William Boecklen – Biology
Julie Fitzsimmons – Art
Pierre Orelus – Curriculum & Instruction
Jane Smith – Nursing
Kassia Wosick – Sociology

September 11, 2017

Overview

The Committee for the Assessment of Student Learning in General Education (CASL-GE) is responsible for assessing student achievement of general education (GE) learning outcomes and for supporting improvement of student learning in lower-division general education courses. Annual assessments, performed at a program level rather than course by course, measure student achievement of general education (GE) learning outcomes related to a subset of New Mexico's mandated general education competencies. During the 2015-2016 academic year, the committee performed an assessment of fine art competencies among undergraduate students primarily ranked as juniors or seniors, students who had presumably completed the majority of their required lower-division GE courses. A pilot assessment of an instrument to assess NM HED cultural awareness in fine arts competencies was also completed and revised with planned use during the 2016-2017 academic year.

Assessment Process

NM HED Area I Communication Competencies (b, c, and d) and Area IV Humanities and Fine Arts Competencies (d)

The cultural awareness in fine arts assignment and scoring rubric were developed by a working subcommittee of the CASL- GE and implemented on a pilot scale during 2015 and 2016. Two pilot sessions were required to finalize instrument questions and wording. In the assignment, students read a brief passage on the topic of the Art in the Public Places Act (AIPP Act) and then were asked to write an essay in response to one of two prompts as to whether the legislature should continue to invest in public artworks. The instructions indicated the essay would be evaluated based on how they organized and expressed their ideas; on their use of rhetorical strategies to persuade, inform, and engage the audience; on the appropriate use of grammar and mechanics; and on their ability to evaluate the cultural environment enrichment aspect to a contemporary issue. These elements correspond to the following NM HED Areas I and IV competencies:

- I(b) – Express a primary purpose in a compelling statement and order supporting points logically and convincingly (Rubric Question I)
- I(c) – Use effective rhetorical strategies to persuade, inform, and engage (Rubric Question II)
- I(d) – Employ writing and/or speaking processes such as planning, collaborating, organizing, composing, revising, and editing to create presentations using correct diction, syntax, grammar, and mechanics. (Rubric Question IV)
- IV(d) – Draw on historical and/or cultural perspectives to evaluate any or all of the following: contemporary problems/issues, contemporary modes of expression, and contemporary thought. (Rubric Question III)

These also align with the Baccalaureate Experience Effective Communication and Academically Prepared outcomes.

Assessment data were collected by CASL-GE members or their representatives from randomly selected Viewing a Wider World (VWW) courses. The majority of students in VWW courses are juniors or seniors, and most have completed all or a majority of their lower-level GE courses.

Following an IRB-approved procedure, the assignment administrator read a script to students regarding the purpose of the assessment and the voluntary nature of their participation. Students desiring to participate were provided with an informed consent letter (requiring their signature) and were given 30 minutes to complete the assignment. Data were collected from a total of 317 students in 7 different VWW courses during Spring Semester 2016.

In May 2016, Assessment Liaisons and interested faculty participated in a scoring session held at the NMSU Teaching Academy. A total of over twenty-five faculty members attended the scoring session. Participants were trained in use of the assignment rubric by Dr. Ryan Goss, chair of the CASL-GE Committee, and Dr. David Smith, director of assessment. Scoring of student work then proceeded according to the following process:

- Four to five graders were seated at each of several tables in the scoring room, with each assigned a grader number.
- A single copy of each student paper was printed and identified only by the student's Banner ID number. Stacks of student papers and rubrics were placed on each of the tables.
- Graders scored individual student papers according to the rubric, recorded scores and their grader number on the rubric sheet, and stapled the sheet to the back of the student paper. Each question was graded according to the following scheme: 0 = no evidence of competence, 1 = emerging, 2 = competent and 3 = skillful.
- Each student paper was scored independently by two different graders. CASL-GE facilitators compared scores and, when differing by more than 1 on any item, initiated a third grader to improve agreement.
- Once a student paper was scored by two graders, it was submitted to an administrative assistant who recorded scores and the students Banner ID in an Excel spreadsheet.

All data from the scoring session were submitted to the Office of Institutional Analysis for processing to determine information on range, central tendency, and inter-rater reliability, and for correlation of those items with student class standing (junior/senior or freshman/sophomore).

An Open Forum at the NMSU Teaching Academy for discussion of the findings of this assessment is planned for fall 2017. The results of will be discussed with the next GE instrument 2016-2017 cycle which focuses on cultural awareness in diversity and economic class.

The Area V cultural awareness in diversity and economic class instrument under development at this time was administered in a similar fashion to over 20 students in a single VWW class in the spring. Student work was scored by CASL-GE committee members, followed by discussion and minor revisions of the instrument and scoring rubric. Results of the pilot assessment will not be discussed further in this document.

Cultural Awareness in Fine Arts Findings

Analysis Pools

Statistical findings were determined for four different student pools:

- All students enrolled in selected VWW courses (N = 317)
- Students with junior or senior standing enrolled in selected VWW courses. (N = 236)

- Students with freshman or sophomore standing enrolled in selected VWW courses. (N = 50)
- Numerous missing student ID or mismatched student ID prevented direct demographic comparisons of results.

Tabulated Findings

Results for all students are listed in Table 1. The data columns correspond to the four rubric categories discussed above. The condensed range category of ≥ 2 corresponds to the % of students scoring as “competent” or “exemplary”. Because there were two to three scores per assessment, and these scores were averaged, the results are reported by range of scores. Students scored highest on Assessment I with 60% in the 2.5-3.0 range. Scores were lower on Assessment III, with 12% in the 0-0.9 range and 28% in the 2.5-3.0 range.

Table 1: Findings for all students

Statistics	Assessment I		Assessment II		Assessment III		Assessment IV		Composite	
	Percent	N	Percent	N	Percent	N	Percent	N	Percent	N
Score Range										
0-0.9	1%	1	5%	9	12%	23	1%	1	1%	2
1.0-1.4	6%	12	18%	34	13%	25	8%	15	10%	19
1.5-1.9	8%	15	16%	30	11%	22	13%	25	25%	49
2.0-2.4	26%	50	24%	47	36%	69	40%	77	30%	57
2.5-3.0	60%	115	38%	73	28%	54	39%	75	34%	66
Total	100%	192	100%	193	100%	193	100%	193	100%	193
Central Tendency										
Mean	2.45		1.95		1.85		2.17		2.11	
Median	2.5		2		2		2		2.13	
Condensed Range										
<2.0	15%	28	38%	73	36%	70	21%	41	36%	70
≥ 2.0	85%	165	62%	120	64%	123	79%	152	64%	123
Total	100%	193	100%	193	100%	193	100%	193	100%	193
Krippendorff's Alpha										
Alpha	0.557		0.571		0.454		0.442		0.66	

When considering competency level (≥ 2.0 as competent), 64% of students scored at or above the competency level on their composite score; 85% of students scored competent or above on Rubric Question I “Responds Directly to the Question;” 62% of students scored competent or above on Rubric Question II “Explains Response with Effective Strategies;” 64% of students scored competent or above on Rubric Question III “Student Addresses/Evaluates Cultural Environment Aspect;” and 79% of students scored competent or above on Rubric Question IV “Student Employs Writing Processes with Correct Syntax, Grammar, and Mechanics.”

Inter-rater reliability (IRR) was measured for each category on the rubric using Krippendorff’s alpha, a numerical measure of the extent to which two or more raters agree with each other when assigning a score to the same artifact. The observed disagreement between evaluators is corrected by the amount of disagreement expected by chance. It ranges from 0.0 (no agreement) to 1.0 (perfect

agreement). Scores above 0.667 indicate an acceptable level of inter-rater reliability.

In general measures of agreement, the following guidelines are given:

- 0.8 and 1 Very good agreement
- 0.6 and 0.79 Good agreement
- 0.4 and 0.59 Moderate agreement
- 0.2 and 0.39 Fair agreement
- 0.0 and 0.19 Poor agreement

For all areas of assessment, the inter-reliability rating was “moderate” and “good” for the composite score.

Representativeness of Sample

A Chi-Square test of homogeneity compared the demographic characteristics of students enrolled in the courses selected for the assessment against students not in the course sample to ensure the students in the selected courses were representative of the broader NMSU population. There were no statistically significant differences between groups with respect to minority status, first generation status, class, or student population. However, the assessment course sample did differ with respect to other categories. These are summarized in Table 2. Because of lack of identifying information on many of the assessments, it was not possible to compare the demographics of the students who took the assessment with the population in the sampled courses nor to the junior/senior population.

Demographic Variable	P-Value	Possible Sample Biases
Citizenship	<.0001	Non-resident aliens over-represented
Ethnicity/Race	0.0003	Non-resident aliens over-represented
Minority	0.3698	None
Gender	0.0019	Males over-represented
First Generation	0.0823	None
Class	0.8163	None
Student Population	0.9816	None
College	<.0001	Business college over-represented
Status (FT vs. PT)	<.0001	Full-time over-represented

Table 2: Sample

Representativeness (Students in VWW Assessment Courses Vs Non-Takers)

Discussion

Student Scores

Close to two-thirds of the student artifacts received competent or higher for each of the scoring categories. Over three-quarters of the student artifacts directly responded to the question and used correct syntax, grammar and mechanics in their writing. These numbers indicate students can successfully respond to the writing prompt and use appropriate writing basics in that response. Writing fundamentals appear to be appropriate for this level of writing assignment for a vast majority of the students.

A smaller majority of students were able to use effective writing strategies to persuade with evidence and were able to address or evaluate the cultural environment and enrichment aspect of the writing prompt. The slightly lower competencies values for these criteria could be attributed to the individual student course selection from a variety of art courses within the current NMSU GE course framework.

Reliability Issues

In previous GE assessment cycles, interrater reliability were poor. However, “Krippendorf’s alpha” inter-rater reliability scores (Table 1) were moderate to good across all questions. The increased efforts to improve reliability of scores by repeating the pilot process, by strengthening the norming process and by encouraging the scores to adopt a more consistent scoring appear to be fruitful for increasing reliability. However, additional measures will be needed in future cycles to consistently improve reliability to good to very good.

Measures of the representativeness of sample (Table 2) showed significant differences between enrolled students in selected VWW courses and the general NMSU population for several demographic categories including the over-representation of males and students from the Business College. This likely reflects similar representations of students enrolled in the particular VWW courses selected for data collection. While the significance of these findings is unclear, they do suggest that greater care may be needed in selection of courses for data collection to assure a more representative sample.

Because of the insufficient student IDs to conduct demographic comparisons, efforts are needed in future annual cycles to ensure the IDs are correct. Students should be able to confirm their student ID number and facilitators should remind students to write down their ID number.