UNIVERSITY OF
TEXAS
ARLINGTON

INSTITUTIONAL EFFECTIVENESS AND REPORTING

ASSESSMENT OF SOCIAL RESPONSIBILITY USING AN ADAPTED AAC&U VALUE RUBRIC AT THE UNIVERSITY OF TEXAS AT ARLINGTON

Summer 2023 Report

Assessment of Social Responsibility Using an Adapted AAC&U Value Rubric

As outlined by the Texas Higher Education Coordinating Board (THECB), the Texas Core Curriculum specifies the knowledge and skill areas required for college education. Social Responsibility is one of the six Texas Core Curriculum Objectives mandated by THECB (2013). Social Responsibility involves an individual's obligation to act in ways that benefit society, show concern for the environment, and balance the welfare of others with self-interest. THECB defines Social Responsibility as "intercultural competence, knowledge of civic responsibility, and the ability to engage effectively in regional, national, and global communities."

UT Arlington assesses the Texas Core Curriculum Objectives through a cyclical rotation plan, measuring six Core Objectives, including Social Responsibility, every three years. This report summarizes the assessment results for Social Responsibility based on papers collected from two sections of a Women and Gender Studies course in the Language, Philosophy, and Culture foundational component area.

Methods

Participants

Written samples of Signature Assignments were collected from 78 undergraduates enrolled in two sections of a Women's Studies (WOMS 2310) course, which is part of the Core Curriculum and offered every semester at UT Arlington. This course attracts a diverse group of students, with a higher enrollment of female students compared to male students.

IER staff reviewed ungraded papers to determine their suitability for inclusion in the scoring session and deidentified them to remove names and other personal information, thereby protecting student privacy. Most student participants (92.31%; n = 72) were female. The racial and ethnic composition of the participants was diverse and reflective of the overall campus population. One-third of the participating students identified as Hispanic/Latino (33.33%, n = 26), and one-fourth identified as Black/African American (24.36%, n = 19). Table 1 presents the summary demographic information of the students whose work is included in this analysis.

Table 1. *Summer 2023 Social Responsibility Student Demographics*

Categorical Information	N	Percentage
Gender		
Female	72	92.31%
Male	6	7.69%
Racial/Ethnic Description		
Asian	7	8.97%
Black, African American	19	24.36%
Foreign, Non-Resident Alien	5	6.41%
Hispanic, All Races	26	33.33%
Multiple Races/Ethnicities	4	5.14%
Not Specified	1	1.28%
White, Caucasian	16	20.51%
Level		
Freshman	4	5.12%

Sophomore	22	28.21%
Junior	28	35.90%
Senior	24	30.77%
First Generation Student		
Yes	33	42.31%
No	45	57.69%
Pell Grant Recipient		
Yes	43	55.13%
No	35	44.87%
UTA Enrollment Year		
2022 – 2023	4	5.13%
2021 – 2022	38	48.72%
2020 – 2021	21	26.92%
2019 – 2020	10	12.82%
2018 – 2019 and prior	5	6.41%
Transfer Student		
Yes	17	21.79%
No	61	78.21%

Students pursuing degrees in nine UT Arlington colleges and schools were represented in the sample data. Majority of participants were from College of Nursing and Health Innovation, College of Liberal Arts, and College of Science (see Table 2).

Table 2: Students by Colleges/Schools

College/School	N	Percentage
College of Architecture, Planning & Public Affairs	3	3.85%
College of Nursing and Health Innovation	28	35.90%
College of Business	2	2.56%
College of Education	2	2.56%
College of Engineering	7	8.97%
College of Liberal Arts	17	21.79%
College of Science	15	19.24%
Division of Student Success	1	1.28%
School of Social Work	3	3.85%

Assessment Instrument

Evidence of Social Responsibility in the Signature Assignment was measured using an adapted rubric, specifically the AAC&U Intercultural Knowledge and Competence Rubric (AAC&U, 2009). The AAC&U VALUE Rubrics were developed as part of a national initiative to assess student learning outcomes in the core curriculum and have been extensively studied and validated for over ten years. In this adapted version, five of the six measures from the original rubric were retained verbatim, while the Verbal and Non-verbal Skills dimension was removed to better align with implementation strategies at UTA. The dimensions included in the present analysis were: 1) Knowledge: Cultural Self-Awareness, 2) Knowledge: Knowledge of

Cultural Worldview Frameworks, 3) Skills: Empathy, 4) Attitudes: Curiosity, and 5) Attitudes: Openness. For details on all dimensions, see Appendix A. The rubric functions as a matrix that provides narrative descriptions of expected work quality and corresponding point values for scoring the six measures. The point values range from 1 to 4, with 1 indicating baseline performance (Benchmark-1), 2 indicating approaching milestone (Milestone-2), 3 indicating achieved milestone (Milestone-3), and 4 indicating the highest mastery (Capstone-4) of Social Responsibility. AAC&U, the authors of the rubric, permit zero ratings if the paper does not meet the minimum content or quality standards defined in the rubric. Numerical ratings were set above the benchmark according to AAC&U recommendations (Greenhoot & Bernstein, 2012). The attainment target was set at a score of 2 (Milestone-2).

Raters, Rater Calibration, and Scoring

Twelve qualified UTA faculty and staff raters with advanced degrees relevant to the core objective being assessed participated in an in-person scoring session on campus in August 2023. During the session, each rater read and rated each paper silently in a group setting. To facilitate tracking, each rater was assigned a unique code number, which was included with the rating sheet to allow IER to monitor which papers were rated by each individual.

An IER staff member, who was qualified to lead the discussion on the core curriculum goals and facilitate the rater calibration process, conducted the scoring day. During calibration, the facilitator guided the raters through a discussion of the rubric dimensions and helped the group operationalize the levels for each skill measure. After discussing the rubric, the facilitator used one student work sample as an anchor paper for the calibration process. All raters scored the anchor paper across all five rubric dimensions, which was followed by a facilitated group discussion to clarify each dimension of the Social Responsibility rubric and to develop a shared approach to scoring.

The scoring process began once the raters felt comfortable with the dimensions and rating intervals. At least two raters reviewed each paper, and ratings were assigned using the rubric. Ratings were collected as they were completed and entered into a spreadsheet by IER staff, who reviewed the data to ensure no missing ratings or other concerns that may need immediate resolution. If the values of the skill measure scores for a paper from the two raters were identical or within one point difference, then the two scores were considered in agreement and averaged. For example, if Rater A scored the Empathy measure with a value of 2 and Rater B scored the same measure with a value of 3, then the rating was considered in agreement, and scores for that dimension were averaged, resulting in a score value of 2.5. However, if there was a difference of more than two points on any single dimension, a third rater was asked to read and assign scores for the paper. In such cases, three scores were averaged together to determine the final score. For example, if Rater A scored the Empathy measure with a value of 1 and Rater B scored the same measure with a value of 4, the rating was not in agreement, and a third rater was asked to read and score the paper.

Analysis and Results

Inter-rater Agreement

To evaluate the reliability of the assessment process, agreement between raters was analyzed to see how frequently the two raters agreed on scoring. The inter-rater agreement was observed throughout the scoring session to determine if re-calibration on one or more scale dimensions was necessary due to frequent low agreement. During the Social Responsibility scoring session, no re-calibration was needed.

The percentage of agreement between raters was calculated to see how

frequently the two raters agreed on scoring for the same student and to gauge the effectiveness of the assessment process. The percentage of agreement among raters for all categories remained between 81% and 91% for the five SR categories (see Table 3). Generally, a minimum of 70% inter-rater agreement is considered a baseline required agreement. Agreement scores above 70% indicate that the scoring is reliable..

Table 3. *Scoring Agreement Percentage Among Raters for SR Skills Dimensions*

Dimension (UTA Social Responsibility VALUE Rubric)	Percentages
Knowledge (Cultural self-awareness)	91%
Knowledge (Knowledge of cultural worldview frameworks)	86%
Skills (Empathy)	86%
Attitude (Curiosity)	81%
Attitude (Openness)	82%

Note: The agreement percentage was computed by dividing the number of agreements by the total number of ratings

Apart from the simple percentage agreements, researchers widely measure the reliability of rating agreements between different raters to eliminate chance agreements using Intraclass Correlation Coefficient, the measure of the consistency among raters when scoring the same subjects independently. In the Social Responsibility scoring process, all raters have advanced degrees and work experience and attended the same training just before the scoring session to mitigate biasness and chance agreemen. The inter-rater agreement was also computed to follow best research practices. The extent to which different raters agree on their judgments establishes the validity and credibility of measurements or ratings.

The inter-rater agreement was determined to check the consistency level of the rating by calculating the Intraclass Correlation Coefficient (ICC). High ICC values indicate more reliability between rater scores. Commonly accepted guidelines were used to interpret the ICC results. These suggest that the range of 0.40 to 0.74 is considered fair to good inter-rater agreement, with results above 0.74 classified as excellent inter-rater agreement and results lower than 0.40 considered poor inter-rater agreement (Fleiss, 1986; Shrout & Fleiss, 1979). ICC values for the Social Responsibility scoring session are presented in Table 4.

Across the five measures, the ICC values showed fair to good agreement levels between raters. The agreement level for three dimensions was above 0.60: Cultural Self-Awareness (ICC = 0.66), Knowledge of Cultural Worldview Frameworks (ICC = 0.61), and Openness (ICC = 0.64). The ICC values for the other two measures, Empathy (ICC = 0.54) and Curiosity (ICC = 0.51) were also above 0.50 (see Table 4). Having these high levels of agreement suggests that applying the rubric to the assignment resulted in sufficient reliability, while training of the raters was adequate for the scoring process, suggesting that the evidence of student attainment can be evaluated confidently.

Table 4. *Intraclass Correlation Coefficient for Social Responsibility Skills Dimensions*

Dimension (SR VALUE Rubric)	Coefficient
Knowledge (Cultural Self-Awareness)	0.66
Knowledge (Knowledge of Cultural Worldview Frameworks)	0.61

Skills (Empathy)	0.54
Attitude (Curiosity)	0.51
Attitude (Openness)	0.64

Note 1: *less than 0.40 = poor agreement; between 0.40 and 0.74 = fair to good agreement; greater than 0.74 = excellent agreement.*

Note 2: *The intra-class correlation coefficient (ICC) was calculated as a two-way random effects model. Values in this model type with random rater pairings are typically expected to be lower than those where rater pairings are fixed throughout the rating day.*

Student Performance

All five dimensions' mean and standard deviation scores were calculated to understand overall student performance. Across the five dimensions, students scored highest in Knowledge (Cultural Self-awareness; mean = 1.79), while Attitude (Openness) had the lowest score (mean = 1.24). A rating of two indicates that the dimension milestone minimum was met. For the analyzed data, means and standard deviations for each dimension of the Social Responsibility rubric are presented in Table 5.

Table 5. *Means for Social Responsibility Skills Measure Scores*

Measurement Dimensions	N	Mean	SD
Knowledge (Cultural Self-Awareness)	78	1.79	0.80
Knowledge (Knowledge of Cultural Worldview Frameworks)	78	1.51	0.85
Skills (Empathy)	78	1.58	0.91
Attitude (Curiosity)	78	1.28	0.92
Attitude (Openness)	78	1.24	0.92

Overall, the average performance of these UT Arlington undergraduate students was above the baseline (Benchmark 1) for all dimensions. The student average scores were between 1.58 and 1.79 for three dimensions: Knowledge (Cultural Self-Awareness), Knowledge (Knowledge of Cultural Worldview Frameworks), and Skills (Empathy). The average score for the remaining two dimensions, Attitude (Curiosity) and Attitude (Openness), remained below 1.50. It is worth noting that the standard deviation scores were high for all dimensions and close to a whole rating interval. For all dimensions other than Knowledge (Cultural Self-Awareness), analysis suggests that zero ratings were assigned with high frequency (as discussed below). In summary, most papers included in the sample were determined to be above Benchmark-1 level.

Observations and Limitation

This report presents results from communal rating sessions of UTA student work to assess Social Responsibility achievement as part of the core curriculum. The assessment of Signature Assignments used an adapted rubric to measure student learning in Social Responsibility domain—a modified version of the AAC&U Intercultural Knowledge and Competence VALUE rubric. We sampled evaluated student papers from two sections of a Women and Gender Studies course within the Language, Philosophy, and Culture foundational area. Results revealed a pattern of strengths and weaknesses in demonstrating social responsibility skills in the evaluated papers and suggested opportunities for

improvement.

The student artifacts analyzed came from students with a diverse range of degree paths. These papers were predominantly written by students in their junior or senior years at UTA. The sample was also highly diverse in terms of race/ethnicity, though more than 90% of participants were female students.

The analysis showed that students achieved scores above Benchmark-1 level but couldn't reach attainment target of Milestone-2. Ratings for the Curiosity and Openness dimensions were noticeably lower, with a higher rate of zero scores assigned. The high number of zero scores indicates that raters observed little to no evidence of curiosity and openness in the papers and that these dimensions were absent from the signature assignment. Future ratings may improve if the signature assignment instructions align more closely with all dimensions of the rubric used in the evaluation and THECB expectations.. This alignment will help increase the likelihood of achieving the desired core curriculum objectives.

Alignment between the Signature Assignments and the VALUE rubrics is essential for reliable ratings of student work. Although instructors determine the composition of the Signature Assignment, it may be beneficial to offer suggestions for aligning these assignments with the rubric to instructors whose courses are included in the core curriculum.

References

Association of American Colleges and Universities (2015). Intercultural Knowledge and Competence VALUE rubric. Retrieved on 08/15/2016 from https://www.aacu.org/value/rubrics/intercultural-knowledge

DiPaolo, T. (2016). Program: Social Responsibility Week - 8-12 AUG 2016. Retrieved on 8/18/2016 from http://leaptx.org/author/tdipaolo/.

Finley, A. P. (2012). How reliable are the VALUE Rubrics? *Peer Review, (14)* 1, 31-33.

Fleiss J. L. (1986). The design and analysis of clinical experiments. New York: John Wiley & Sons.

Greenhoot, A. and Bernstein, D. (2011). Using VALUE rubrics to evaluate collaborative course design. *Peer Review 13/14(1).*

Hart Research Associates. (2013). *It Takes More Than a Major: Employer Priorities for College Learning and Student Success.* Washington, DC: Association of American Colleges and Universities.

Hart Research Associates. (2010). *Raising the Bar: Employers' Views on College Learning in the Wake of the Economic Downturn.* Washington, DC: Association of American Colleges and Universities.

Hart Research Associates. (2008). How Should Colleges Assess and Improve Student Learning? Employers' Views on the Accountability Challenge. Washington, DC: Association of American Colleges and Universities.

Lederman, D. (2015). Are they learning? New effort aims to standardize faculty-driven review of student work. *Inside Higher Ed.* Retrieved on 11/13/2015 from https://www.insidehighered.com/news/2015/09/25/new-effort-aims-standardize-faculty- driven-review-student-work.

Shrout, P., & Fleiss, J. L. (1979). Intraclass correlation: uses in assessing rater reliability. *Psychological Bulletin, 86(2),* 420--428.

Appendix A: Social Responsibility VALUE Rubric

SOCIAL RESPONSIBILITY RUBRIC

	Capstone	Milestones		Benchmark
	4	3	2	1
Knowledge *Cultural self- awareness*	Articulates insights into own cultural rules and biases (e.g. seeking complexity; aware of how her/his experiences have shaped these rules, and how to recognize and respond to cultural biases, resulting in a shift in self-description.)	Recognizes new perspectives about own cultural rules and biases (e.g. not looking for sameness; comfortable with the complexities that new perspectives offer.)	Identifies own cultural rules and biases (e.g. with a strong preference for those rules shared with own cultural group and seeks the same in others.)	Shows minimal awareness of own cultural rules and biases (even those shared with own cultural group(s)) (e.g. uncomfortable with identifying possible cultural differences with others.)
Knowledge *Knowledge of cultural worldview frameworks*	Demonstrates sophisticated understanding of the complexity of elements important to members of another culture in relation to its history, values, politics, communication styles, economy, or beliefs and practices.	Demonstrates adequate understanding of the complexity of elements important to members of another culture in relation to its history, values, politics, communication styles, economy, or beliefs and practices.	Demonstrates partial understanding of the complexity of elements important to members of another culture in relation to its history, values, politics, communication styles, economy, or beliefs and practices.	Demonstrates surface understanding of the complexity of elements important to members of another culture in relation to its history, values, politics, communication styles, economy, or beliefs and practices.
Skills *Empathy*	Interprets intercultural experience from the perspectives of own and more than one worldview and demonstrates ability to act in a supportive manner that recognizes the feelings of another cultural group.	Recognizes intellectual and emotional dimensions of more than one worldview and sometimes uses more than one worldview in interactions.	Identifies components of other cultural perspectives but responds in all situations with own worldview.	Views the experience of others but does so through own cultural worldview.
Attitudes *Curiosity*	Asks complex questions about other cultures, seeks out and articulates answers to these questions that reflect multiple cultural perspectives.	Asks deeper questions about other cultures and seeks out answers to these questions.	Asks simple or surface questions about other cultures.	States minimal interest in learning more about other cultures.
Attitudes *Openness*	Initiates and develops interactions with culturally different others. Suspends judgment in valuing her/his interactions with culturally different others.	Begins to initiate and develop interactions with culturally different others. Begins to suspend judgment in valuing her/his interactions with culturally different others.	Expresses openness to most, if not all, interactions with culturally different others. Has difficulty suspending any judgment in her/his interactions with culturally different others, and is aware of own judgment and expresses a willingness to change.	Receptive to interacting with culturally different others. Has difficulty suspending any judgment in her/his interactions with culturally different others, but is unaware of own judgment.

UNIVERSITY OF TEXAS ARLINGTON

AAC&U Association of American Colleges and Universities

Adapted for the University of Texas at Arlington
from AAC&U's Intercultural Knowledge VALUE Rubric
Last Revised 08/01/2016