



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

SPRING **2016** REPORT

Assessment of Social Responsibility Using an Adapted AAC&U Value Rubric at
The University of Texas at Arlington

Institutions of higher education in the United States currently offer students a wide array of choices in terms of academic courses and extracurricular activities. Courses and activities aim to help students to effectively learn, live, and work alongside others. Ideas about an individual's duty to act in ways that benefit society are often included in definitions of social responsibility. These definitions also mention accountability to others, concern for the environment, and balancing the welfare of others versus self. Whether the process of learning social responsibility is taught in mandatory Texas Core Curriculum courses or experienced in university-sponsored programs on campus (e.g., student affairs programming, residential hall activities, group learning sessions such as MAV1000, athletics, etc.), these experiences are an important part of the mission of The University of Texas at Arlington (UT Arlington). Indeed, developing social responsibility is reflected in our mission statement.

"The University of Texas at Arlington is a comprehensive research, teaching, and public service institution whose mission is the advancement of knowledge and the pursuit of excellence. The University is committed to the promotion of lifelong learning through its academic and continuing education programs and to the formation of good citizenship through its community service learning programs. The diverse student body shares a wide range of cultural values and the University community fosters unity of purpose and cultivates mutual respect. "

As designed by the Texas Higher Education Coordinating Board (THECB), the Texas Core Curriculum safeguards the value of undergraduate education in the state by specifying certain knowledge and skill areas which must be included. For example, *Social Responsibility* is one of the

six Texas Core Curriculum Objectives required by the Texas Higher Education Coordinating Board (THECB, 2013). THECB defines *Social Responsibility* as "intercultural competence, knowledge of civic responsibility, and the ability to engage effectively in regional, national, and global communities."

Employers of college graduates agree that *Social Responsibility* is an essential skill. Surveys of employers (Hart Research Associates, 2013, 2010, 2008) reveal that intercultural skills and the ability to respond to new situations are characteristics that hiring managers find desirable. The survey results note that beyond content knowledge, the college experience needs to build a student's sense of *Social Responsibility* (e.g., intercultural understanding and civic involvement).

Developing a written assignment that assesses evidence of a student's *Social Responsibility* presents quite a challenge. In fact, a state-wide initiative recently conducted a week-long Social Responsibility event that was focused on opening a dialog among faculty to identify and refine assignments for this purpose ([DiPaolo, 2016](#)). In the present study, the presence of *Social Responsibility* was measured in student essays using a well-vetted rubric developed by the Association of American Colleges and Universities ([AAC&U](#); Rhodes, 2010).

UT Arlington assesses the Texas Core Curriculum Objective by following a cyclical rotation plan whereby the complete set of six Core Objectives are measured every three years. Requirements, defined by THECB, vary for each of the eight Foundational Component Areas (FCA), making the multi-year cycle ideal. The assessment of *Social Responsibility* is required within five FCAs: Language, Philosophy and Culture; Creative Arts; American History; Government/Political Science; and Social and Behavioral Sciences. This report summarizes *Social Responsibility* data collected from two FCAs (e.g., Creative Arts; Social and Behavioral Sciences). In addition, a separate report is available from the Government/Political FCA, a UT Arlington department that conducts assessment using an individual department method.

Method

Participants

Written samples of the Signature Assignments from courses that were designated as part of the Core Curriculum at UT Arlington were collected from enrolled undergraduates ($N = 301$). Each written essay was rated by a team of trained faculty, separate from the grading process conducted by the instructor of the class. Half of the student participants were female (56%; $n = 169$). The racial and ethnic composition of the participants was predominantly non-Caucasian and thus characteristic of the overall campus population, the fifth most diverse four-year university in the nation (see Table 1). As expected in a sample of Core Curriculum classes, student enrollment consisted of primarily freshman and sophomores. Nine UT Arlington colleges and schools were represented (see Table 2).

Table 1
Student Demographics

Categorical Information	Number of Students	Percent
Gender		
Female	169	56
Male	129	43
Missing or unspecified	3	1
Ethnic Description		
American Indian or Alaskan Native	1	<1
Asian	35	12
Black, African American	42	15
Foreign, Non-Resident Alien	8	3
Hispanic, All races	90	30
Multiple Ethnicities	15	5
Unknown or Not Specified	4	1
White, Caucasian	106	35
Level		
Freshman	170	57
Sophomore	73	24
Junior	28	9
Senior	29	9
Unknown	1	<1

Table 2

Students by College/School

College/School	Number of Students	Percent
College of Architecture, Planning & Public Affairs	4	1
College of Business	18	6
College of Education	5	2
College of Engineering	31	10
College of Liberal Arts	44	15
College of Nursing and Health Innovation	23	8
College of Science	40	13
School of Social Work	38	13
University College or Major Intended	98	33

Procedure

Undergraduate students who were enrolled in sections of Creative Arts and Social and Behavioral Sciences courses during the spring 2016 semester completed a Signature Assignment approved by the UT Arlington Core Curriculum Committee. The Signature Assignment in Creative Arts directed the students ($n = 221$) to write an essay that described their response to the performance of a play that included familiar *Social Responsibility* themes: **1**) cultural self-awareness, 2) stereotypes, and 3) the meaning of race. The other FCA used an ethnography assignment in which the students ($n = 80$) wrote an essay after conducting an interview with someone who had come to the United States from another country. After completion, student Signature Assignments were collected from the department. The papers were assigned a tracking number and then any personal identification information (e.g., the student's name) was removed in preparation for Scoring Day. Two papers (from the analysis of play essays) could not be rated because the copy that was submitted had one or more missing pages. These essays were dropped from the sample, resulting in set of 299 essays to rate.

Assessment Instrument

Evidence of *Social Responsibility* within the Signature Assignment was measured using an adapted rubric (e.g., AAC&U Intercultural Knowledge and Competence Rubric; AAC&U, 2015;

see Figure 1). Of the six measures from the Intercultural Knowledge and Competence Rubric, five were retained verbatim from the original rubric. The adapted rubric simply dropped the *Verbal and non-verbal skill* area. A team of faculty representing institutions across the United States developed this rubric for use across academic disciplines, as part of a national initiative called Valid Assessment of Learning in Undergraduate Education (VALUE; AAC&U, 2015). The adapted rubric is organized into five measures: 1) Knowledge: *Cultural Self-awareness*, 2) Knowledge: *Knowledge of cultural worldview frameworks*, 3) Skills: *Empathy*, 4) Attitudes: *Curiosity*, and 5) Attitudes: *Openness*. The rubric contains a matrix that provides a narrative description of the expected quality of work and corresponding point values for scoring. The point values ranged from 1 - 4 with 4 representing the highest mastery of *Social Responsibility*. Raters assigned a score for each measure.

SOCIAL RESPONSIBILITY RUBRIC

	Capstone 4	Milestones		Benchmark 1
		3	2	
Knowledge <i>Cultural self-awareness</i>	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 <i>Knowledge of cultural worldview frameworks</i>	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 <i>Empathy</i>	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 <i>Curiosity</i>	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 <i>Openness</i>	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.

Figure 1. UT Arlington Social Responsibility Rubric

Attainment targets (numerical ratings) were set in concordance with recommendations gathered from AAC&U research (Greenhoot & Bernstein, 2012). As such, the attainment target was set at a value of two (Milestone) which reflects that the student demonstrated skills above the Benchmark level.

Raters, Rater Calibration, and Scoring

We recruited raters for scoring the Signature Assignments from among the UT Arlington academic community. The ratings were completed in a group setting on Scoring Day. Assignment of a code number to each rater allowed us to track the papers rated by each person and ensured their anonymity in the final dataset. Using the rubric, student work samples were rated by faculty and staff with advanced degrees. A majority of the faculty were female (70%, $n = 14$), Caucasian (90%, $n = 18$), and had earned a Ph.D. (55%, $n = 11$). On average the group had 10 years of teaching at the university level and represented the College of Business, College of Education, College of Liberal Arts, College of Nursing and Health Innovation, the Office of the Provost, and the English Language Institute. Several had earned certifications that placed them at high levels within their respective fields, such as Registered Nurse, Texas teaching certificate (K-12), certificate in teaching English as a second language (TESOL), certificate in mediation, certificate in online teaching, and certified community-based literacy instructor.

Two steps were followed before rating the student essays. First, the raters gathered on Scoring Day and completed a facilitated rater-calibration process. For example, after listening to a facilitator present the rubric and operationalize the levels for each skill measure, two samples of student work were used as anchor papers in the hands-on rater calibration process. During this step, based on the five skill measures of the rubric, all of the raters read and scored the anchor papers to

identify exemplars. Afterwards, the facilitator led a group discussion aimed at reaching a common understanding of *Social Responsibility*.

Next, the actual scoring process began. Each paper was assigned to a minimum of two raters and each one scored the paper using the rubric. If the values of the skill measure scores for a paper from the two raters were identical or within two points, then the two scores were averaged. For example, if Rater A scored the *Cultural Self-awareness* measure with a value of 2 and Rater B scored *Cultural Self-awareness* with a value of 4, then the score for that dimension was averaged, resulting in a score value of 3. If the scores from the two raters differed by more than two points, then a third rater was assigned the paper. In that case, the two most similar scores were averaged together and the third was dropped. Figure 2 contains the rater score sheet.

	Rater Number: <input type="text"/> <input type="text"/>				Rater Number: <input type="text"/> <input type="text"/>			
	4	3	2	1	4	3	2	1
<u>Knowledge</u> : Cultural self-awareness	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<u>Knowledge</u> of cultural worldview frameworks	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<u>Skills</u> : Empathy	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<u>Attitudes</u> : Curiosity	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<u>Attitudes</u> : Openness	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Figure 2. Rater Score Sheet

Analysis and Results

Inter-rater Agreement

In order to evaluate the effectiveness of the assessment process, inter-rater agreement analysis was conducted to see how frequently the two raters agreed on scoring. The inter-rater agreement level was determined by calculating the intra-class correlation coefficient (ICC). High

ICC values indicate more agreement 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 .40 considered poor inter-rater agreement (Fleiss, 1986; Shrout & Fleiss, 1979).

Across the five measures, the ICC values were exceptional. Three were classified as excellent inter-rater agreement: *Cultural Self-awareness* (ICC= 0.77), *Knowledge of cultural worldview frameworks* (ICC= 0.75), and *Openness* (ICC= 0.78). The ICC values for the other two measures, *Empathy* (ICC= 0.725), *Curiosity* (ICC= 0.727), were in the highest range of fair to good inter-rater agreement. Having these high levels of agreement suggests that the application of the rubric to the assignment was accomplished with sufficient reliability and that the evidence of student attainment of the skills can be evaluated with confidence.

Student Attainment

Frequencies were used to examine students' *Social Responsibility* across five measures (see Table 3). Scores from both Rater1 and Rater2 were counted, resulting in a total of 598 scores/row, in which each row represents the skill measure across the 299 student papers. Mean scores exceeded targets suggested by the AAC&U.

Summary

The current assessment of Signature Assignments used a rubric adapted from the AAC&U Intercultural Knowledge and Competence VALUE rubric. Results revealed a pattern of strengths and weaknesses for a sample of undergraduate students. The student scores were strongest for the dimensions of *Empathy* and *Openness*. That said, all five measures met attainment targets. In four of the five skills measures, ten percent or less of the students in this group of UT Arlington undergraduates fell below Milestone 2. It is also important to note that at the other end of the spectrum, a similar percentage of students surpassed both Milestone levels (e.g., 2 and 3) and

Table 3

Skill Measure scores for Social Responsibility from the adapted Intercultural Knowledge and Competence VALUE Rubric

Skill Measure	Mean (SD)	Score Frequency (n, %)			
		4	3	2	1
<i>Cultural Self-awareness</i>	2.54 (0.71)	51 (9%)	281 (47%)	206(34%)	60 (10%)
<i>Knowledge of cultural worldview frameworks</i>	2.53 (0.68)	53 (9%)	258 (43%)	241 (40%)	46 (8%)
<i>Empathy</i>	2.69 (0.65)	71(12%)	297 (50%)	203 (34%)	27 (5%)
<i>Curiosity</i>	2.36 (0.73)	46 (8%)	207 (35%)	261 (44%)	84 (14%)
<i>Openness</i>	2.67 (0.74)	74 (12%)	306 (51%)	162 (27%)	56 (9%)

reached the Capstone 4 level. In that sense, a standard normal curve was observed for the rubric with a single mode at the Milestone 3 level for all measures except *Curiosity*.

Inter-rater agreement for rating the essays using the rubric was very positive. Calibration activities to operationalize the levels of attainment described in the rubric followed recommended reliability guidelines (Finley, 2012). In addition, the use of anchor papers as a basis for discussing examples of each level was positively received by the cadre of faculty raters. In that sense, this study provides additional support for the use of calibration activities to increase rating reliability prior to commencing with reading and rating the student essays (Finley, 2012).

While, the sample was representative of the rich diversity of the campus population at UT Arlington, sample essays were drawn from two of five FCAs and that could have limited the scope of the summary report. That said, students represented all nine schools and colleges. In addition, because, in concept, all students regardless of their major must take forty-two hours of approved Texas Core Curriculum courses, each FCA should contain a representative sample of the academic

community at large (as demonstrated in Tables 1 and 2). It is nevertheless recommended that future studies should sample *Social Responsibility* in the three other FCAs.

A strength of the study was sampling from two different types of assignments, especially in light of the challenge of creating assignments that ask students to respond to each of the skill measures. Because of the naturalistic design of the assessment plan for collecting Signature Assignments, the essays from some departments did not contain enough elements of *Social Responsibility* to rate them using the rubric. Further, though both of the assignments in the sample covered five of the six measures well, evidence of verbal and non-verbal communication was not present and the university dropped it in its adapted rubric. In addition, between the two Signature Assignments, the interview assignment seemed to engage students in expressing their *Curiosity* and *Openness* about other cultures more than the play critique essay. The assignment prompt for the interview essay explicitly asked the students to express curiosity about another culture both to uncover surprises they encountered and to discuss their own reactions. The two assignments seemed to cover the other measures without significant differences.

Overall, this initial assessment of *Social Responsibility*, a THECB Core Objective, was very positive. On average, the student scores met and in many cases, exceeded attainment targets. The high levels of inter-rater agreement suggest that the rater calibration activities were helpful and that the scores obtained are reliable measures of *Social Responsibility*. Future studies will continue to examine trends in student performance related to these measures among undergraduate students at UT Arlington.

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>.
- Rhodes, T. (Ed.). (2010). *Assessing outcomes and improving achievement: Tips and tools for using rubrics*. Washington, DC: Association of American Colleges and Universities.
- Shrout, P., & Fleiss, J. L. (1979). Intraclass correlation: uses in assessing rater reliability. *Psychological Bulletin*, 86(2), 420--428.