



Cardinal Core Pilot Diversity Assessment (Spring 2022)

Cardinal Core Program

The Cardinal Core program at the University of Louisville prepares students to do the advanced work needed for their baccalaureate degrees and prepares them to contribute to society throughout their lives through their professional work and civic engagement. The program emphasizes the development of key intellectual skills relevant to any career path: critical thinking, quantitative reasoning, effective communication, and the understanding of historical, social, and cultural diversity. Students will develop these intellectual skills in the following content areas of Arts and Humanities, Historical Perspectives, Oral Communication, Quantitative Reasoning, Social and Behavioral Sciences, Natural Sciences, Written Communication, and the competency area of Diversity in the United States and Globally. Upon completion of the program, students will be prepared to analyze complex problems and evaluate possible courses of action in an environment characterized by diversity and the need for sustainable solutions.

Assessment Administration

The assessment of student learning outcomes is a national expectation in higher education. Section 8.2.b of the Southern Association of Colleges and Schools' (SACS) accreditation standards requires that the institution identify student learning outcomes for collegiate-level general education competencies in its undergraduate degree programs, assess the extent to which it achieves these outcomes, and provide evidence of seeking improvement based on analysis of the results. Further, the Kentucky Council on Postsecondary Education (CPE) states that "All Kentucky public universities and KCTCS colleges are expected to assess, in accordance with SACS-COC Principles of Accreditation and based upon nationally accepted standards, the student learning outcomes associated with their general education programs, indicate a relationship to the faculty-generated Statewide General Education Student Learning Outcomes, and provide evidence of ongoing assessment that ensures comparability for transfer purposes on a three-year cycle."

The Cardinal Core Curriculum Committee (CCCC) is charged with continued oversight of the assessment of student learning outcomes across the Cardinal Core curriculum to support the continuous improvement of the Cardinal Core program in alignment with SACS and CPE requirements. The assessment operates on a three-year cycle, in which samples of student work are collected from one content area each semester and assessed by a panel of trained faculty. The CCCC began a pilot of the Association of American Colleges and Universities (AAC&U) VALUE (Valid Assessment of Learning in Undergraduate Education) Rubrics in the first cycle of the Cardinal Core Assessment. Specifically, the Critical Thinking, Intercultural Knowledge and Competence, Oral Communication, Quantitative Literacy, and Written Communication VALUE Rubrics were used to measure the Cardinal Core program's overarching intellectual

skills of critical thinking, effective communication, quantitative reasoning, and social, historical, and cultural diversity.

Through the pilot of the AAC&U VALUE rubrics, the CCCC determined that the Intercultural Knowledge and Competence VALUE Rubric did not provide a reliable and consistent measure of the Cardinal Core program’s overarching diversity competency across all courses. The rubric was designed to focus more on global diversity issues and had limitations in measuring diversity in the localized setting.

As a result, the CCCC developed a new Diversity rubric, adapted from the AAC&U Intercultural Knowledge and Competence VALUE Rubric and the University of Western Oregon Diversity Rubric. A pilot of the new instrument was conducted in Spring 2022. The pilot assessment was to ensure that the rubric will work across all Cardinal Core courses and also provide some baseline performance data prior to entering the next three-year cycle of assessment in Fall 2022.

DRAFT Cardinal Diversity Rubric

The Draft Cardinal Core Diversity Rubric follows the same format as the AAC&U VALUE Rubrics, as well as the UofL Cardinal Core Natural Sciences Rubric. The rubric uses four scoring categories, with 4 indicating performance of the measure as “capstone” level, 3 indicating performance at “milestone,” 2 indicating “milestone,” and 1 indicating performance at “benchmark.” In addition, a score of zero can be assigned to any work that does not meet the level 1 performance. The University of Louisville further disaggregates the zero option into a “not applicable” rating that can be selected for assignments that did not provide an opportunity for the student to demonstrate the criterion within the rubric measure, as opposed to the student simply not demonstrating the rubric criteria.

The rubric is provided in Figure 1. The competency names for each row were considered draft and reviewers were encouraged to focus more on the language in the performance cells rather than the competency name. For the purposes of the pilot, the competencies were given generic headers of Row 1-4.

		4	3	2	1
Row 1	Knowledge of social, historical, and cultural structures¹	Demonstrates sophisticated understanding of the complexity of elements relevant to what people do in different contexts in relation to culture, history, economy, politics, communication styles, beliefs, values, or practices, etc.	Demonstrates adequate understanding of the complexity of elements relevant to what people do in different contexts in relation to culture, history, economy, politics, communication styles, beliefs, values, or practices, etc.	Demonstrates partial understanding of the complexity of elements relevant to what people do in different contexts in relation to culture, history, economy, politics, communication styles, beliefs, values, or practices, etc.	Demonstrates superficial understanding of the complexity of elements relevant to what people do in different contexts in relation to culture, history, economy, politics, communication styles, beliefs, values, or practices, etc.

¹ Adapted from the AAC&U’s Intercultural Knowledge and Competence VALUE Rubric and Western Oregon University Diversity Rubric

Row 2	Applied Knowledge²	Applies understanding of multiple perspectives, experiences, and/or power structures to evaluate relevant local and/or global issues or topics, with appropriate depth, breadth, and synthesis.	Applies understanding of multiple perspectives, experiences, and/or power structures to adequately evaluate local and/or global issues or topics.	Applies understanding of multiple perspectives, experiences, and/or power structures to partially evaluate local and/or global issues or topics.	Simplistically applies understanding of multiple perspectives, experiences, and/or power structures when evaluating local and/or global issues or topics.
Row 3	Cultural Self-awareness³	Articulates complex insights into cultural identity (e.g., norms, biases) in local and/or global contexts.	Recognizes new perspectives about cultural identity (e.g., norms, biases) in local and/or global contexts.	Identifies some aspects of cultural identity (e.g., norms, biases) in local and/or global contexts.	Shows minimal awareness of cultural identity (e.g., norms, biases).
Row 4	Attitudes/Openness⁴	Demonstrates cultural self-awareness when considering new ways of being and thinking in the world.	Begins to demonstrate cultural self-awareness when considering new ways of being and thinking in the world.	Has difficulty considering new ways of being and thinking in the world.	Unable to set aside own biases and assumptions when considering new ways of being and thinking in the world.

**Edited by the Cardinal Core Curriculum Committee (CCCC) 11.12.21*

Figure 1. Draft Cardinal Core Diversity Rubric

Assessment Process

For the Spring 2022 pilot of the new Diversity rubric, student work was collected from across the Arts & Humanities, Social & Behavioral Sciences, Historical, Perspectives, and Written Communication content areas. Each academic unit that offers these courses was included in the assessment sample and the sample represented courses with and without Diversity competency designations (D1 and D2). While U.S. Diversity (D1) and Global Diversity (D2) courses have specific Diversity student learning outcomes, it is an expectation of all Cardinal Courses to address the overarching competency of understanding of historical, social, and cultural diversity.

The Cardinal Core Office collaborated with department chairs and faculty teaching the courses selected for the pilot to collect syllabi, assignment prompts, and student work samples. After the semester withdrawal deadline passed, the Cardinal Core Office retrieved the class rosters for all of the selected courses from the Office of the Registrar and selected a random sample of students from each of the courses. Instructors were sent the list of students selected for the assessment along with detailed instructions requesting that instructors provide a copy of one assignment prompt along with the ungraded responses for the selected students to be sent via email to the Cardinal Core Office service account.

² Adapted from the AAC&U's Intercultural Knowledge and Competence VALUE Rubric and Western Oregon University Diversity Rubric

³ Adapted from the AAC&U's Intercultural Knowledge and Competence VALUE Rubric and Western Oregon University Diversity Rubric

⁴ Adapted from the AAC&U's Intercultural Knowledge and Competence VALUE Rubric

Student artifacts were collected and stored in an electronic repository and uploaded into the LiveText© assessment management system. A panel of faculty (tenured and tenure-track faculty, term faculty, and adjunct faculty) assessed student artifacts.

Prior to the assessment reading, assessors completed an online training module featuring an overview of the assessment process and detailed dissection of the assessment rubric. Assessors were asked to complete four benchmark sample assessments prior to attending a synchronous norming session via Microsoft Teams. Benchmarks were assignments selected to represent a wide range of content and skill development to give the assessors a baseline for measuring expectations of learning and evaluating student performance (Herman, Osmundson, & Dietel, 2010). During the norming session, faculty engaged in discussion about the benchmark assessment scores to share their rationales for why particular scores were selected. The results from scoring benchmark samples are provided in Appendix A.

After completion of the training and rubric norming, each faculty assessor was assigned a username and password for one of three LiveText© accounts and a list of courses and sections to assess. Three readers assessed each artifact so that scores could be compared across assessors for reliability purposes.

Data Collection Overview

The sample for the pilot of the new Diversity rubric included courses from the Arts & Humanities, Social & Behavioral Sciences, Historical, Perspectives, and Written Communication content areas. Table 1 provides a breakdown of the number of work samples by Cardinal Core designation. The samples came from courses in the College of Arts & Sciences, College of Education and Human Development, Kent School of Social Work, and the School of Public Health.

Table 1

Breakdown of Sample by Cardinal Core Designation

Cardinal Core Designation	Number of Work Samples
AH	28
AHD1	25
AHD2	27
SB	31
SBD1	28
SBD2	26
SBH	10
SBHD1	10
SBHD2	21
WC	12

Summary of Assessment Data

For the pilot assessment, 218 student artifacts from 51 courses were assessed. Table 2 provides the percentage of work samples scored at each rubric level for the Diversity rubric. The rubric provided a “0” score and for purposes of understanding why a “0” was assigned, the Cardinal Core Assessment provided reviewers with the option of “Not Applicable”. The “Not Applicable” indicates an absence of the assessment criteria due to the type of assignment, while a “0” indicates that the student could have demonstrated the criteria and did not.

Table 2

Percentage of Artifacts Scored at Each Rubric Level for Diversity Rubric (Three Reviews)

	Row 1	Row 2	Row 3	Row 4
Capstone (4)	8.5%	9.0%	8.2%	11.6%
Milestone (3)	31.5%	25.8%	24.9%	42.4%
Milestone (2)	28.1%	26.3%	33.6%	16.2%
Benchmark (1)	17.1%	23.8%	16.0%	4.3%
(0)	5.8%	5.8%	5.5%	7.6%
Not Applicable (NA)	9.0%	9.3%	11.8%	17.9%

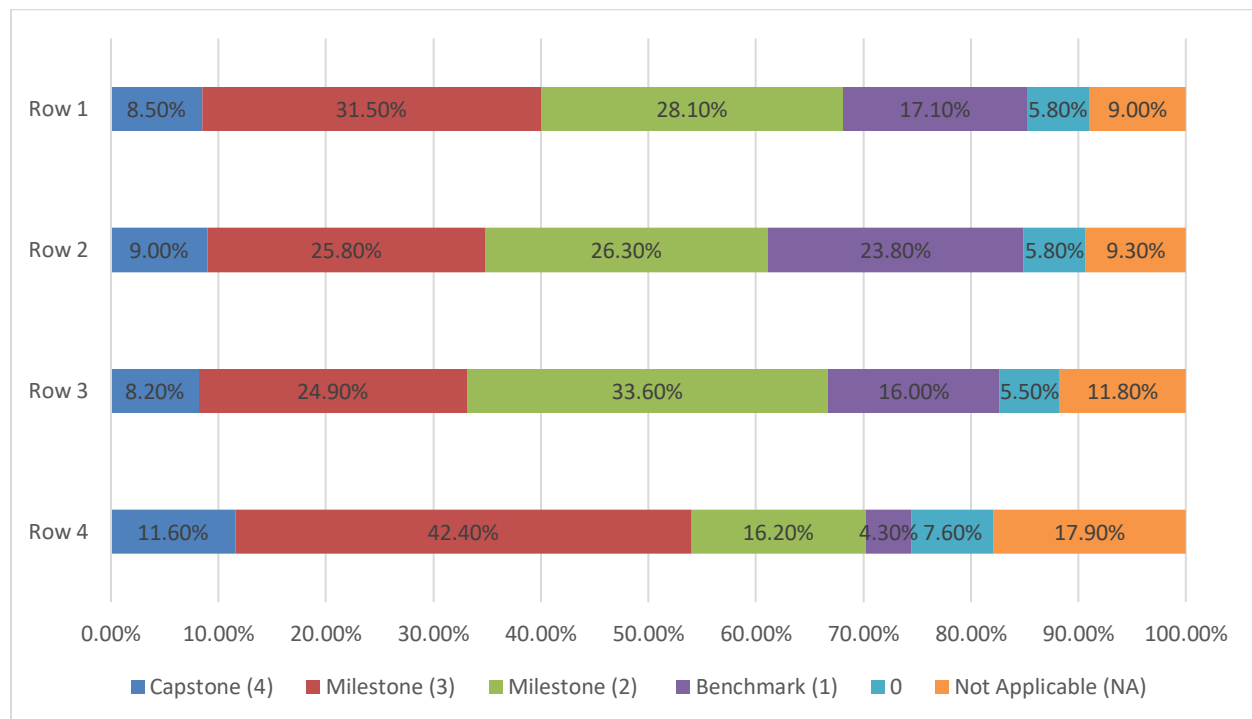


Figure 2. Percentage of Artifacts Scored at Each Rubric Level for Diversity Rubric (Three Reviews)

In alignment with AAC&U VALUE Institute practices, scores for each individual work sample were calculated based upon scores assigned by three separate reviewers. The scores from all three reviewers were averaged and rounded to determine individual work sample scores for each rubric row. “Not Applicable” ratings were treated as “0”s to determine a work sample score.

Table 2 provides the percentage of work samples scored at each level with scores averaged (one score per work sample).

Table 3

Percentage of Artifacts Scored at Each Rubric Level for Diversity Rubric (Scores Averaged)

	Row 1	Row 2	Row 3	Row 4
Capstone (4)	1.1%	2.3%	3.8%	3.6%
Milestone (3)	26.4%	27.3%	30.8%	32.7%
Milestone (2)	24.1%	29.5%	32.1%	28.2%
Benchmark (1)	16.1%	36.4%	25.6%	25.5%
(0)	32.2%	4.5%	7.7%	10.0%

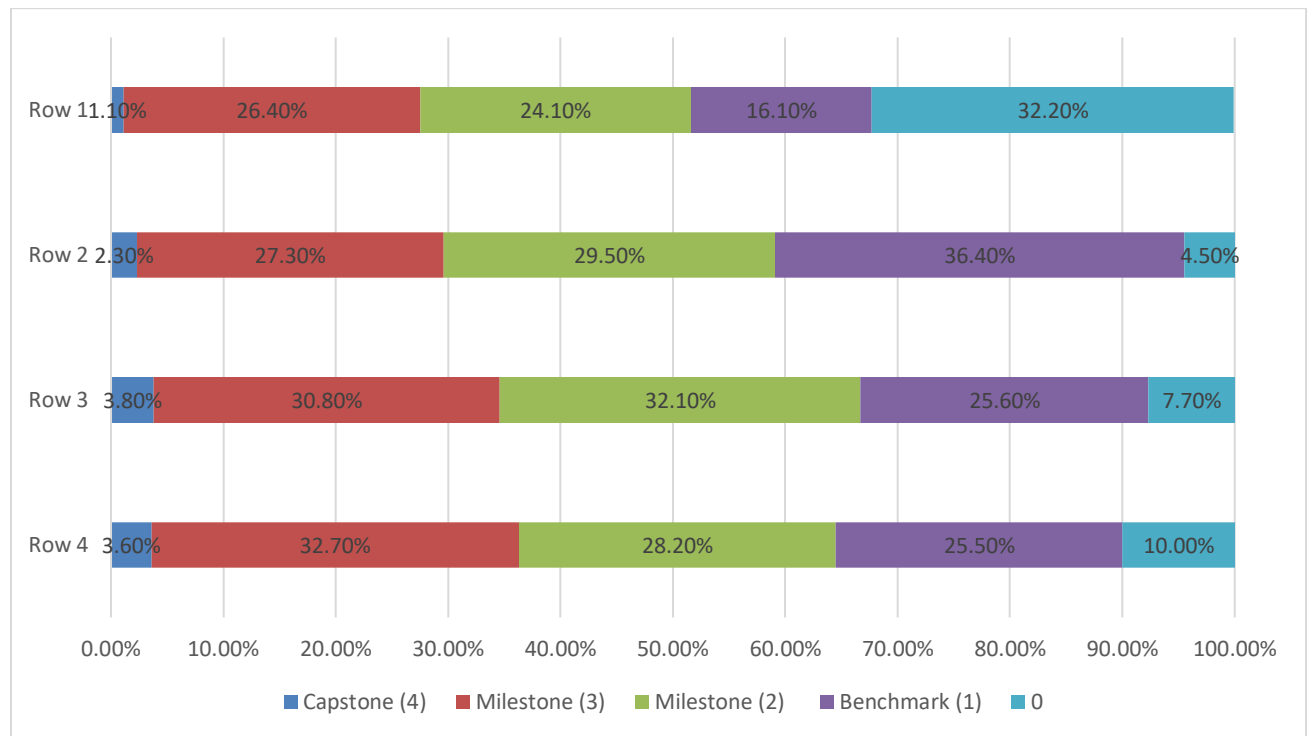


Figure 3. Percentage of Artifacts Scored at Each Rubric Level for Diversity Rubric (Scores Averaged)

The mode for the individual work sample scores is provided in Table 4.

Table 4

Mode of Individual Work Sample Scores

	Mode
Row 1	3
Row 2	1
Row 3	2
Row 4	3

Since the pilot included samples from across the Arts & Humanities, Social & Behavioral Sciences, Historical Perspectives, and Written Communication content areas and were from courses that also fulfilled U.S. Diversity (D1) and Global Diversity (D2) requirements, a comparison of ratings based on diversity classification is provided in Table 5 and Figure 3. Prior to conducting this analysis, individual ratings were calculated for each student work sample. These ratings were calculated by averaging the scores assigned by each of the three reviewers and rounding to the nearest whole number. The individual ratings by rubric category (row) were then compared across groups. The mean for each rubric row for courses without a Diversity designation, courses with a U.S. Diversity (D1) designation, and courses with a Global Diversity (D2) designation is provided in Table 5 and Figure 3.

Table 5

Comparison of Courses by Diversity Designation

	No Diversity Designation	U.S. Diversity (D1) Designation	Global Diversity (D2) Designation
Row 1	1.57	2.16	2.23
Row 2	1.64	2.11	2.00
Row 3	1.61	2.08	2.08
Row 4	1.86	2.30	2.24

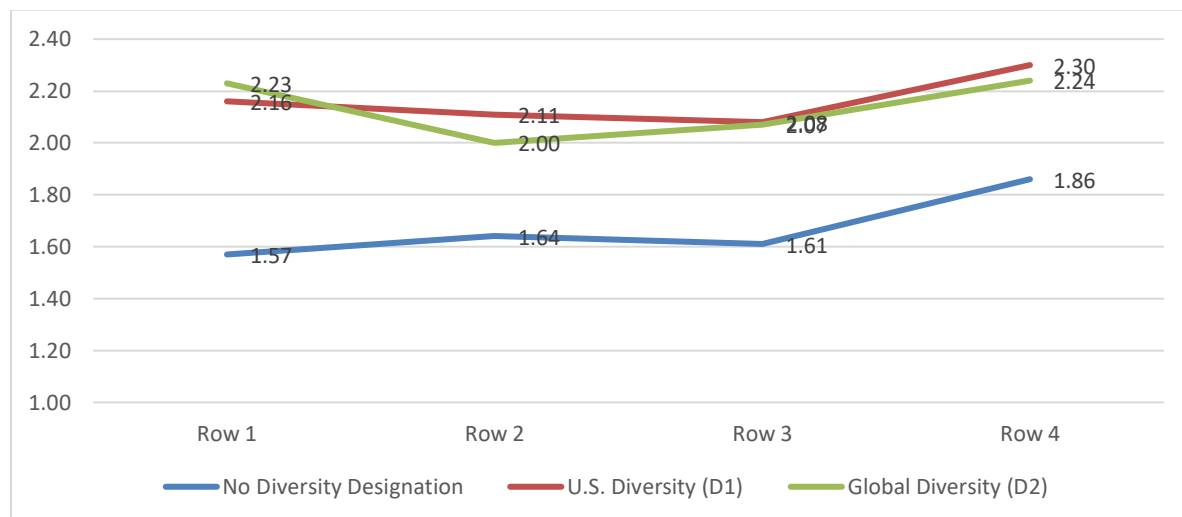


Figure 4. Comparison of Courses by Diversity Designation

A one-way analysis of variance (ANOVA) revealed that there were significant differences in the means of work samples for courses without a Diversity designation, courses with a U.S. Diversity (D1) designation, and courses with a Global Diversity (D2) design work samples on all rubric measures.

There was a significant difference between groups for Row1 [$F(2,200) = 8.34, p < .05$], Row2 [$F(2,203) = 4.36, p < .05$], Row3 [$F(2,203) = 5.99, p < .05$], and Row4 [$F(2,203) = 4.01, p < .05$]. Post hoc comparisons using Tukey HSD test indicated that the mean score for work samples from courses with no Diversity designation was significantly different from courses with a U.S. Diversity (D1) designation on all rubric measures (rows), the mean score for work samples from courses with no Diversity designation was significantly different from courses with a Global Diversity (D2) designation on all rubric measures (rows), and there were no significant differences between samples from courses with a D1 or a D2 designation for any of the rubric measures (rows). The results indicate that students in courses with a D1 or D2 designation scored significantly higher than students in a course without a Diversity designation on all Diversity rubric measures.

Inter-rater Reliability

Three separate readers assessed each student artifact. Table 5 displays the mean score for the three separate readings of all artifacts.

Table 5 provides multiple measures of inter-rater reliability. The percentage agreement value was calculated to determine the percentage of artifacts for which all three assessors scored at the either the same or within one performance level. Values for *Total Agreement* provided in Table 6 represent the percentage of artifacts for which all three assessors selected the same score (e.g., Assessors 1, 2, and 3 all selected 3). *Agreement (within 1 level)* represents the percentage of artifacts for which all three assessors scored the artifact at the same performance level or within one level (e.g., Assessor 1 selected a score of 3, Assessor 2 selected a score of 2, and Assessor 3 also selected a score of 2). If the assessor assigned “not applicable” for the artifact that was treated as a 0 for the inter-rater reliability analysis since a 0 and “not applicable” would both indicate the reviewer did not see the student demonstrate any component of the rubric measure.

In addition to percentage agreement, a one-way, average-measures intra-class correlation coefficient (ICC) was calculated to assess inter-rater reliability. ICC coefficients between .75 and 1.00 are considered excellent, .60 to .74 considered good, .40 to .59 fair, and below .4 is considered poor (Cicchetti, 1994).

Table 6

Inter-rater Reliability for Pilot UofL Cardinal Core Diversity Rubric Measures

Competency Measure	Total Agreement	Agreement (within 1 level)	ICC	95% Confidence Interval
Row 1	12.8%	60.6%	.65	(.57-.73)

Row 2	13.3%	52.3%	.64	(.54-.71)
Row 3	14.2%	55.5%	.62	(.53-.70)
Row 4	17.9%	48.6%	.59	(.48-.67)

Syllabus Review

Consistent with all Cardinal Core content area assessments, syllabi were collected and reviewed for the Cardinal Core student learning outcomes and corresponding methods of assessment.

The review of syllabi sought to answer two questions:

- 1) Does the syllabus contain the content specific Cardinal Core learning outcomes approved for the course?
- 2) Are assessment methods stated that support the content-specific Cardinal Core learning outcomes approved for the course?

Of the 51 courses included in the pilot sample 43 courses (84.3%) had the Cardinal Core learning outcomes in the syllabus and 41 courses (80.4%) had the corresponding methods of assessment. Through the assessment process, faculty that did not have the outcomes and assessments in their syllabus were reminded of the importance of including this information going forward and were directed to the student learning outcomes on the Cardinal Core website.

The Cardinal Core Curriculum Committee (CCCC) has continued to emphasize the importance of incorporating the Cardinal Core learning outcomes into course syllabi. Integration of the Cardinal Core outcomes into the syllabus is one indication to the committee that faculty are incorporating the learning outcomes into the course curriculum.

Summary of Diversity Pilot Findings

Based on the findings from the pilot assessment and feedback from reviewers, the new Diversity rubric is broad enough to apply in both localized and global contexts. There were no significant differences in performance between U.S. (D1) and Global (D2) Diversity courses as compared to the results from the previous assessments with the AAC&U Intercultural Knowledge and Competence VALUE Rubric. While the results indicate that samples from courses without a Diversity designation performed lower than samples from courses with a Diversity designation, this is primarily due to assignments that did not ask students to demonstrate the competencies being measured. Reviewers did express concerns that some of the work samples provided did not ask students to demonstrate understanding of historical, social, and cultural diversity, but there were no concerns raised about applying the instrument in courses that did ask students to demonstrate this in localized versus global settings.

Overall, the reviewers had positive feedback about the assessment instrument. The reviewers suggested that Row 3's competency name be changed to "Cultural awareness" and remove "self" to reflect more accurately what the row is measuring and what students will be asked to demonstrate in their Cardinal Core courses. In addition, some reviewers commented that they rarely assigned anything lower than a "3" rating for Row 4. While this may indicate that UofL students generally exhibit openness to new ways of being and thinking in the world, the committee may want to revisit this row after additional assessments to determine if the criteria are providing an adequate understanding of students' attitudes and openness.

The CCCC voted to approve removing the word "self" from the category name for row 3 of the rubric and proceed with the instrument for the next three-year cycle of assessment. The committee, along with the assessment subcommittee will continue to monitor the fit of the instrument as it is applied across each of the content area assessments.

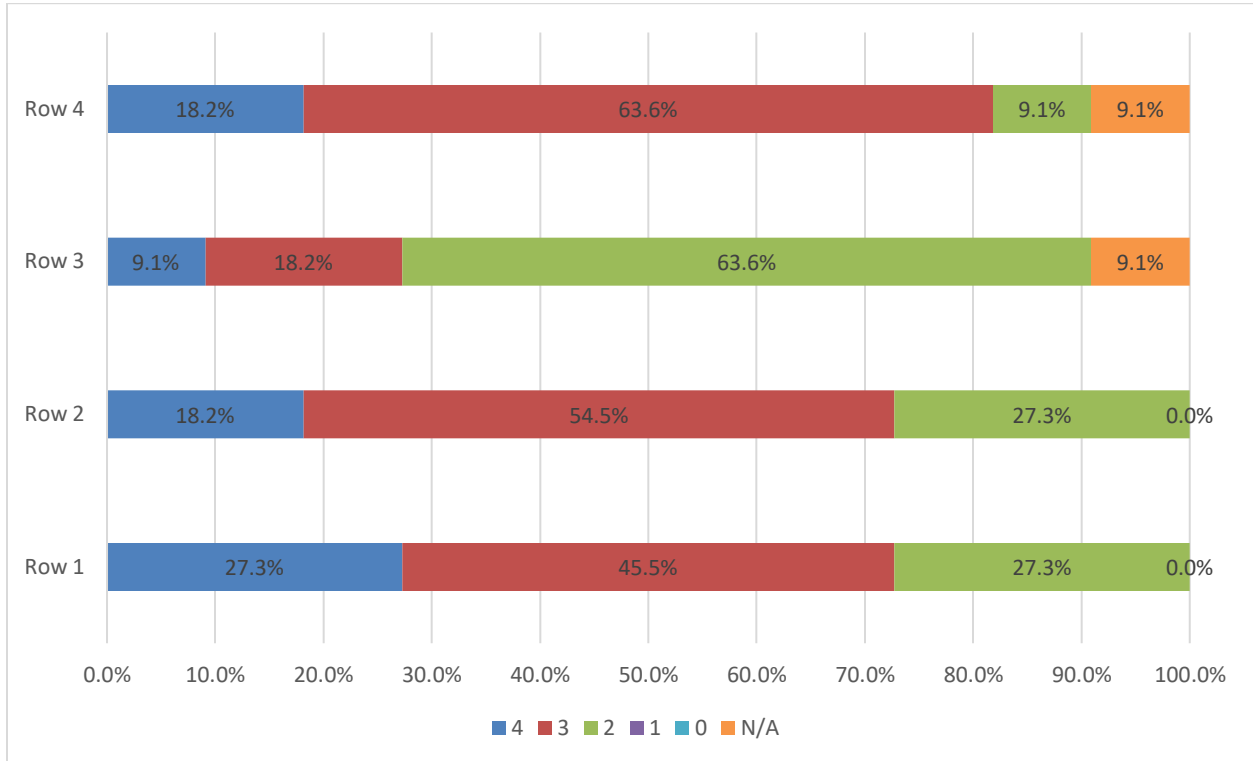
References

- Association of American Colleges and Universities. (2009). *Critical Thinking VALUE Rubric*. Retrieved from <https://www.aacu.org/value/rubrics/critical-thinking>
- Association of American Colleges and Universities. (2009). *Intercultural Knowledge and Competence VALUE Rubric*. Retrieved from <https://www.aacu.org/value/rubrics/intercultural-knowledge>
- Association of American Colleges and Universities. (2009). *Written Communication VALUE Rubric*. Retrieved from <https://www.aacu.org/value/rubrics/written-communication>
- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, 6(4), 284-290. doi: 10.1037/1040-3590.6.4.284
- Fletcher, I., Mazzi, M., & Nuebling, M. (2011). When coders are reliable: The application of three measures to assess inter-rater reliability/agreement with doctor–patient communication data coded with the VR-CoDES. *Patient Education and Counseling*, 82(3), 341-345. doi:10.1016/j.pec.2011.01.004
- Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in Quantitative Methods for Psychology*, 8(1), 23-34.
- Herman, J. L., Osmundson, E., & Dietel, R. (2010). *Benchmark Assessment for Improved Learning*. (AACC Report). Los Angeles, CA: University of California.

Appendix A. Results from Benchmark Samples in the training

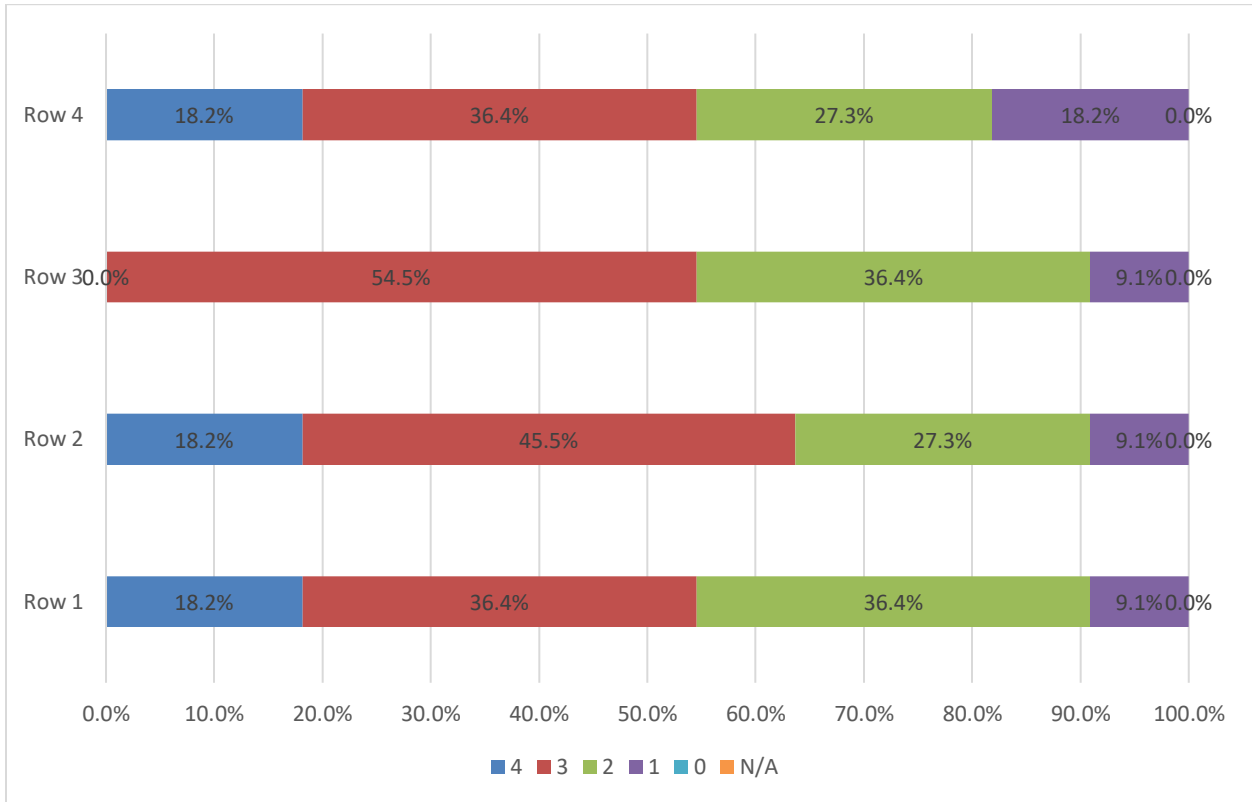
Sample 1

	4	3	2	1	0	NA
Row 1	27.3%	45.5%	27.3%	0.0%	0.0%	0.0%
Row 2	18.2%	54.5%	27.3%	0.0%	0.0%	0.0%
Row 3	9.1%	18.2%	63.6%	0.0%	0.0%	9.1%
Row 4	18.2%	63.6%	9.1%	0.0%	0.0%	9.1%



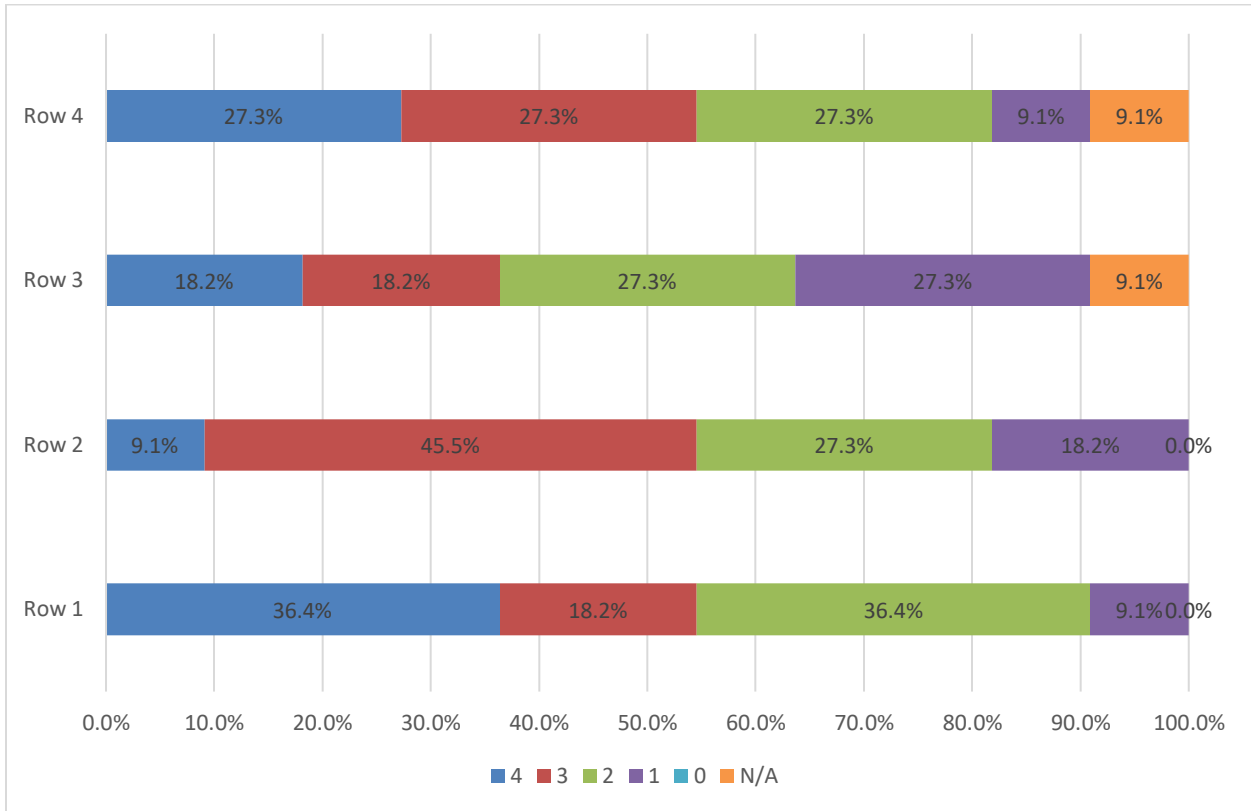
Sample 2

	4	3	2	1	0	NA
Row 1	18.2%	36.4%	36.4%	9.1%	0.0%	0.0%
Row 2	18.2%	45.5%	27.3%	9.1%	0.0%	0.0%
Row 3	0.0%	54.5%	36.4%	9.1%	0.0%	0.0%
Row 4	18.2%	36.4%	27.3%	18.2%	0.0%	0.0%



Sample 3

	4	3	2	1	0	NA
Row 1	36.4%	18.2%	36.4%	9.1%	0.0%	0.0%
Row 2	9.1%	45.5%	27.3%	18.2%	0.0%	0.0%
Row 3	18.2%	18.2%	27.3%	27.3%	0.0%	9.1%
Row 4	27.3%	27.3%	27.3%	9.1%	0.0%	9.1%



*The reviewers engaged in extensive discussion about “cultural awareness” versus “cultural self-awareness” during the synchronous norming session. It was determined that a student can demonstrate cultural awareness without talking about their own personal awareness. The awareness is presented in how they approach and discuss the subjects they are addressing in their work. As a result of the discussion, a poll was conducted to rescore Row 3 and Row 4 for work Sample 3. The results of the poll are provided below.

***Rescored Row 3 and 4 for Sample 3**

	4	3	2	1	0	NA
Row 3	18.0%	54.0%	27.0%	0.0%	0.0%	0.0%
Row 4	0.0%	77.0%	22.0%	0.0%	0.0%	0.0%

Sample 4

	4	3	2	1	0	NA
Row 1	9.1%	63.6%	18.2%	9.1%	0.0%	0.0%
Row 2	0.0%	54.5%	18.2%	27.3%	0.0%	0.0%
Row 3	9.1%	45.5%	36.4%	9.1%	0.0%	0.0%
Row 4	18.2%	63.6%	18.2%	0.0%	0.0%	0.0%

