

# General Education Assessment Report (2021-2022): Diversity



Prepared by: Jacob M. Held, Ph.D.

Assistant Provost for Academic Assessment

and General Education

September 2022

## **Table of Contents**

- I. Prefatory**
- II. Summary**
- III. Diversity**
- IV. Results**
  - A. Goal A (Own)
  - B. Goal B (Other)
  - C. Goal C (Creative Works)
- V. Recommendations**
- VI. Appendix**
  - A. ICC calculations via SPSS
  - B. Scorer comments in AQUA

## **I. Prefatory**

The following report covers assessment efforts for the Diversity Core competency over the academic year of 2021-2022. This report includes a narrative account of all components of the assessment cycle as well as selected data informing those narratives. The intention of this report and the recommendations herein is to be advisory to the UCA Core Council and all relevant stakeholders as stewards of the general education program at UCA.

## II. Summary

The UCA Core is assessed on a four-year cycle. Each year one competency area is addressed. For AY 2021-2022, Diversity was the area scheduled to be assessed.

Assessment in higher education ought to be driven by the idea that reliable data can be used to inform curricular changes to improve student learning. The focus is always on student performance. If you want to improve learning, you must know where your students are and whether or not your curriculum is impactful. Thus, there must be moments of assessment where student performance is measured consistently, according to an objective standard, and across time.

During the AY 21-22 assessment cycle observations were made regarding student performance as well as the process of assessment itself as regards the Diversity competency of the UCA Core. Below are several key takeaways.

- Faculty participation improved during this cycle. **AY 21-22 survey response rate = 63.98%**
- Poorly chosen or designed assignments were a problem frequently noted by the score teams (See Appendix B). Pre-cycle training needs to focus on assignment design and needs to be readily accessible and more widely used by faculty.
- With respect to student learning: **significant growth** was noted in some areas.
- Although the data for Goal A was troubling, for Goal B we saw a decent percentage of students at the upper division score **“accomplished” or higher**. For Goal C, in general, the number of students so scoring was **significant**.
- The **rubrics ought to be revisited** in order to clarify language and allow scorers to develop and impose a standard set of expectations.

The following report provides a detailed presentation and analysis of the assessment process and results for the Diversity competency of the UCA Core during AY 21-22. This report provides an initial interpretation of selected data of the assessment of the Diversity competency.

### III. Diversity

The UCA Core is assessed on a four-year cycle. Each year one competency area is addressed. For AY 2021-2022, Diversity was the area scheduled to be assessed. The semester prior to the academic year scheduled for assessment training sessions were offered for all faculty scheduled to be teaching a course in the Diversity area during AY 21-22. Multiple sessions were scheduled with times scattered throughout the week to offer several opportunities for faculty to attend. Dr. Held facilitated all sessions. Topics included rubric interpretation, assignment design and selection, as well as a briefing on the process of artifact collection and scoring. Sessions were held on-line on April 13, 16, and 22 of 2021.

Overall, attendance at these sessions was poor. Poor attendance at these sessions may translate into poor participation later on. In addition, poor attendance at these information sessions may foment confusion during the academic year during which assessment takes place. Faculty may be unprepared to participate leading to greater work on behalf of assessment staff and poorer results. Since these sessions cover the assessment process as well as assignment design, poor attendance may mean that the artifacts that are received may be poorly designed to register student performance across the learning outcomes the rubrics were designed to measure. In fact, poorly aligned artifacts is a common issue reported by the scoring teams when assessing student work. During spring 2021, the pre-cycle training was conducted on-line in webinar format, recorded, and posted online. Having the materials posted and accessible is also an added benefit to faculty who may wish to review them at their own pace or a more convenient time. After pre-assessment training, the office of assessment prepared to collect student artifacts during AY 21-22.

#### Artifact Collection:

During AY 21-22, the Office of Assessment attempted to collect artifacts from all courses under the Diversity Core competency including all Lower and Upper division courses so designated as well as capstone courses. Faculty teaching these courses were identified through ARGOS. All identified faculty were contacted multiple times by means of email. Faculty were provided with a link to a google form. The form asked for information regarding what artifact would be chosen, when it would be administered to students, and when and how it would be delivered to the Office of Assessment.

#### Survey Response Rate:

	# of unique instructors teaching courses	# of unique instructors responding to survey	% response rate <sup>1</sup>
Fall 2021	107	72	<b>67.29</b>
Spring 2022	54	38	<b>70.37</b>
Summer 2022	25	9	<b>36.00</b>

---

<sup>1</sup> Response rate for the survey. This would not reflect faculty that participated by submitting artifacts but did not complete the survey.

Response rate indicates percentage of respondents in relation to total number of faculty identified as teaching a course identified as assessing under the Diversity competency.

The total response rate was an improvement over the previous year. In fact, it is one of the highest response rates to date. Unfortunately, the survey response rate did not result in a large pool of artifacts. Anomalies exist, such as the fact that although we had 35 discrete faculty indicate that they would be providing artifacts under Goal A during AY 21-22, we only received artifacts for 19 discrete courses under that goal. Goals B and C did not show this kind of disparity. However, it does indicate a breakdown in the collection process.<sup>2</sup>

#### **Review of Artifacts:**

Evaluation of the artifacts took place in August of 2022. The evaluation team was recruited from faculty who had participated in the assessment process. The evaluation team consisted of:

- Rubric A (Own)
  - Lynn Burley (LLLC)
  - S. Lynne Rich (Sociology, Anthropology, Criminology)
- Rubric B (Other)
  - Roger Pauly (History)
  - Carla Gilbreath (Health Sciences)
- Rubric C (Creative Works)
  - Kyle Mattson (School of Communication)
  - Ramón Escamilla (LLLC)

Evaluators were remunerated \$250 per day. During the three day sessions evaluators participated in calibration exercises as well as artifact scoring. Days consisted of routine evaluation work from 8:00 am until 4:30 pm with intermittent breaks as evaluators deemed appropriate.

	# of artifacts available	# of artifacts processed <sup>3</sup>	% of artifacts processed
Goal A	311	183	58.84
Goal B	900	652	72.44
Goal C	613	479	78.14

---

<sup>2</sup> In order to better understand the breakdown of this process, in subsequent years the Office of Assessment will collect more data on response rates. (e.g. actual artifact delivery rates from each CRN and track delivery rates by faculty so chairs can be informed.)

<sup>3</sup> Disregards the number of artifacts receiving a second score. Approximately 20% of artifacts in each goal received a second score in order to calculate inter-rater reliability.

Total	1824	1314	72.04
-------	------	------	-------

### Reliability:

The score teams spent the first half of their first day together engaged in norming exercises. The teams reviewed the rubric and proceeded to evaluate anchor assignments. After each assignment is evaluated the team discussed the results and then proceeded to the next assignment. By the close of the calibration exercise the teams expressed a shared understanding of the rubric and shared expectations. Teams also conferred regularly during scoring to continuously “re-calibrate.”

These norming exercises are intended to insure that regardless of team member the score an artifact receives is consistent. If scorer expectations are consistent, then the data will be consistent and generalizable. Calibration is crucial to reliable data, that is, data that reflects the nature of the artifact, in this case student performance, and not the idiosyncrasies of the scorer.

### Percent Agreement and Interrater Reliability

	% agreement	% disagree at 1 pt. <sup>4</sup>	Intraclass Correlation Coefficient (ICC) <sup>5</sup>	Reliability <sup>6</sup>
Goal A	63.37	86.00	.751	Good
Goal B	44.05	89.66	.286	Poor
Goal C	61.34	87.43	.814	Good

Although percent agreement is not often accepted as a reliable statistic when judging inter-rater reliability, of note in this case is the fact that when scorers did disagree, well over 85% of the time that disagreement was only one point in variance. That indicates that even when scorers disagreed it was minor, indicating a slight disagreement in student performance, not a major incongruity between scorer expectations. This would suggest that the data reliably reflect trends. Using a more standard measure of inter-rater reliability for ordinal values, ICC, we find “good” reliability in the teams scoring goals A and C, and “poor” reliability in the team for goal B. The high measure of reliability in goals A and C is notable. The poorer reliability in Goal B may be attributable to the faculty scorers being from dissimilar academic disciplines, but calibration should have corrected for this. We did have recurring conversations

<sup>4</sup> When scorers did disagree, this is the percent of disagreements between a single level, for example, scorer A = 1, scorer B = 2, or scorer A = 3 and scorer B = 4.

<sup>5</sup> See Appendix A for tables

<sup>6</sup> Based on Terry K. Koo and Mae Y. Li. A Guideline for Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research, *Journal of Chiropractic Medicine*, June 2016, vol. 15(2): 155-63.

throughout the process so the disagreement is anomalous. Regardless, reliability among the teams was in general decent, so the data should reflect the relative student performance on the rubric.

#### **IV. Results<sup>7</sup>**

Assessment in higher education ought to be driven by the idea that reliable data can be used to inform curricular changes to improve student learning and assist faculty in developing pedagogies that are more effective. The focus is always on student performance, the goal is learning.

When interpreting assessment data in higher education it is important to note several points. Firstly, the methodology used is often derived from the behavioral and social sciences. However, the higher education environment makes it difficult, if not impossible, to maintain the conditions necessary for reliable statistical analysis using these methods. Samples are small, or in isolated communities, there are myriad factors influencing any variable, most of which cannot be controlled for, nor is it possible to offer control groups as withholding educational opportunities from students for experimental purposes is unethical. The data collected, therefore, must be interpreted in light of these structural barriers, which are endemic to the nature of the study. But while these barriers cannot be removed, they can be ameliorated.

We can get reliable data in terms of identifying trends so long as we know wherein the problems lie and work intentionally to mitigate them. With Core assessment, we have striven to lessen these barriers where possible. We collect student work from the entire population in order to derive a representative sample. Artifacts are all scored on the same rubric, by a single team of calibrated, trained, faculty scorers, thus increasing interrater reliability. We offer training to faculty on assignment design prior to artifact collection, thus allowing faculty to use individual assignments, not standardized ones, while maintaining a consistency of expectation.

If a general education program is to be assessed for common student learning outcomes at a university the size of UCA, the means by which we are doing so addresses, as well as can be addressed, the limitations inherent in assessment in higher education.

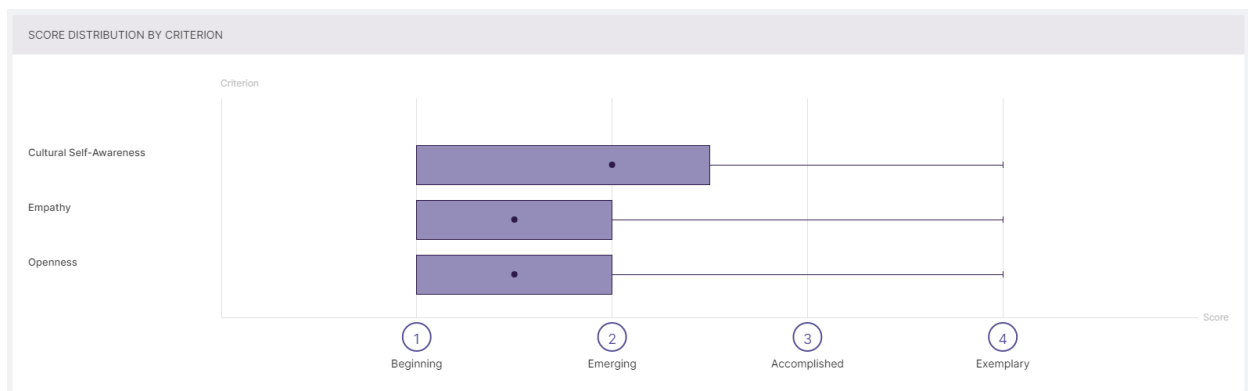
---

<sup>7</sup> Due to an oversight when the AQUA project for this assessment cycle was created, there was no N/A option for when an artifact failed to address a learning outcome. Instead, faculty had to score those artifacts as a “1.” The data was then cleaned up after the fact to remove, when possible, these “1” scores from true “1” scores reflecting actual student performance. (This process was imperfect, so it adds a layer of doubt to the data provided.) This issue has been corrected for future projects.

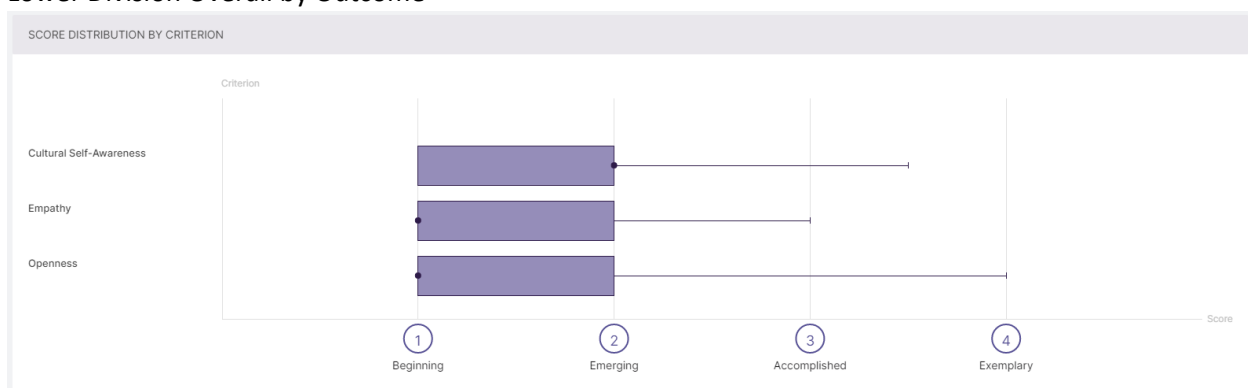
## Goal A (Own)

Goal A presented us with the smallest pool of artifacts. This category is also the smallest category of the Diversity competency area in terms of courses that fall under the goal. 51.25 % of artifacts were scored over the three-day evaluation period.

### Overall Score Distribution by Outcome



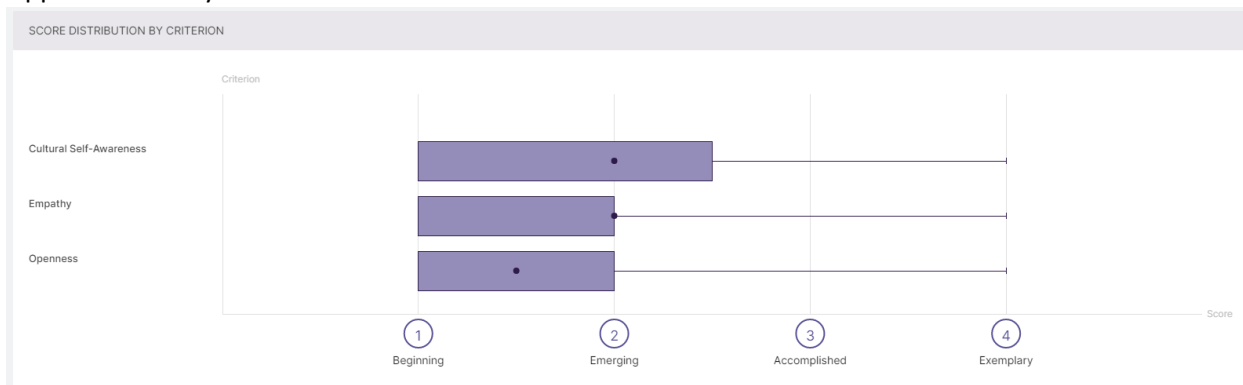
### Lower Division Overall by Outcome



	Max Score	Min Score	Median Score	# of submissions
Cultural Self-Awareness	4	1	2	57
Empathy	3	1	1	57
Openness	4	1	1	57

Averages are not particularly informative for this type of assessment, insofar as we are gauging non-discrete variables, like learning. In addition, at the lower level we would expect to see students at the emerging or developing (1 or 2) levels and that is in fact what we see. Frequency data is much more informative, and then at the upper levels, since it will be more instructive of how well our students have developed across the learning outcomes of the goal. In addition, due to poor faculty participation, and a small category, the LD Core data encompasses only two courses, and so is not statistically reliable, since it reflects those two courses ability to teach across the outcomes more so than a generalizable presentation of student ability as a whole. Thus, these data need to be taken lightly.

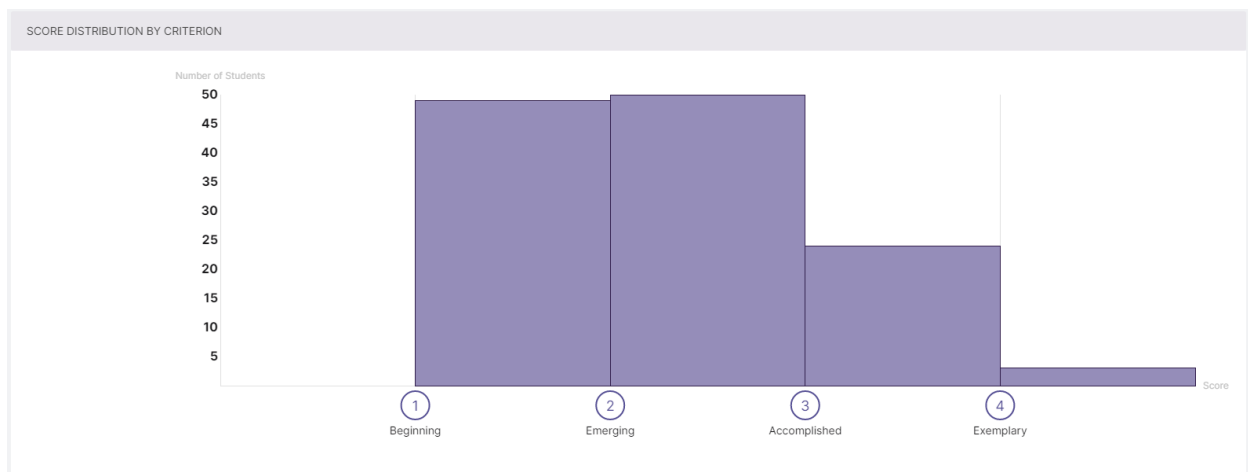
#### Upper Division by Outcome



	Max Score	Min Score	Median Score	# of submissions
Cultural Self-Awareness	4	1	2	126
Empathy	4	1	2	126
Openness	4	1	2	126

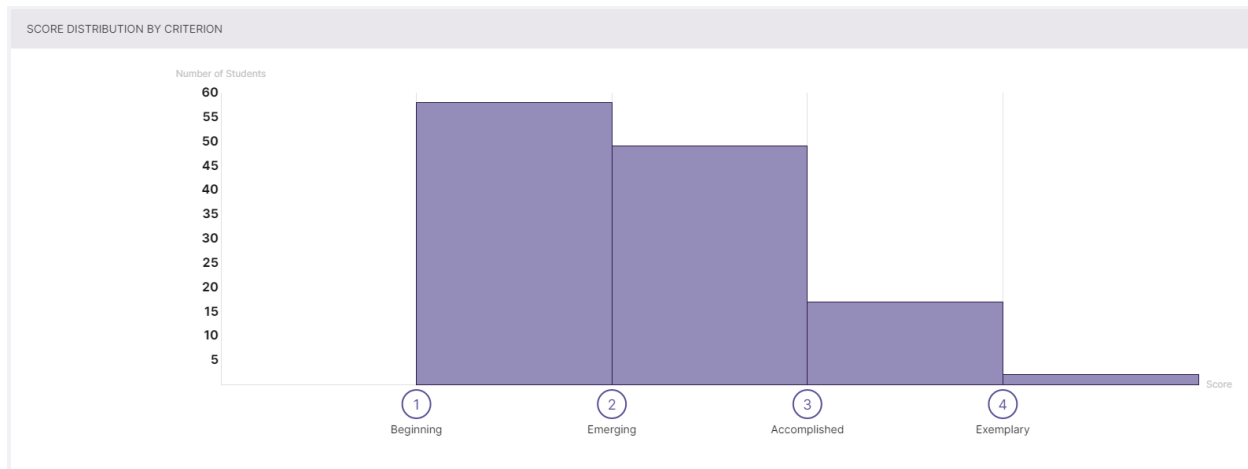
Overall, the upper division scores for Diversity Goal A are troubling. One would expect, or hope, that at the upper division, students would be performing much higher on the rubric, with a significant proportion of students scoring at the 3 (Accomplished) or 4 (Exemplary) levels. Instead, the data looks similar to the LD Core data indicating that students do not progress to greater degrees of competence in this area. Although the maximum score for “Empathy” increases for the upper division courses to a 4, this is clearly an outlier as the median score moves up simply to 2. The same pattern occurs with the “Openness” outcome. The median moves slightly, but not enough to demonstrate marked improvement among students, as a detailed look at the upper division by outcome demonstrates. The frequency data tells the story of student performance in upper division coursework.

## Outcome 1: Cultural Self-Awareness (UD Core)



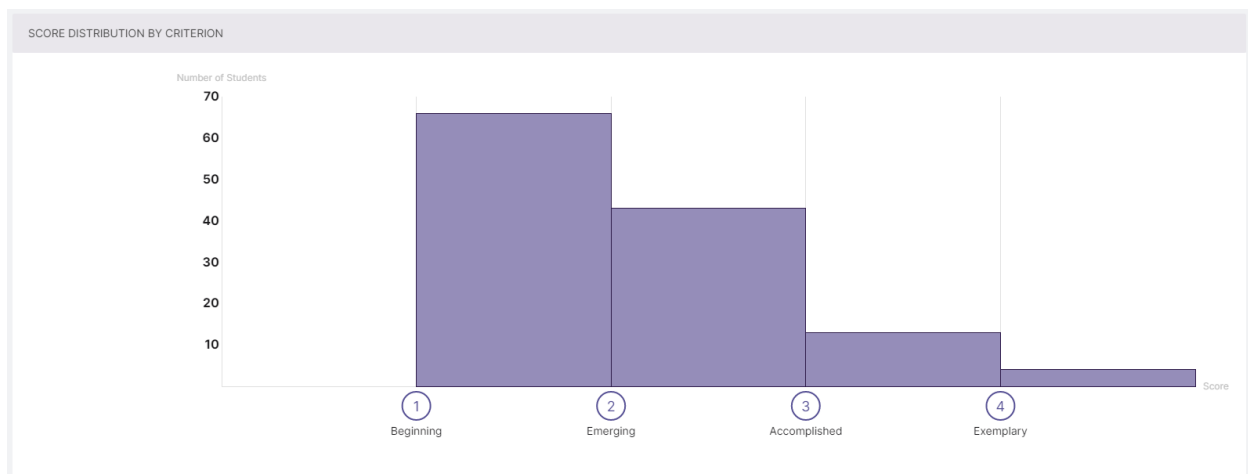
Of note, only 3 of a total 126 artifacts scored registered at level 4, or exemplary. Thus, even though these artifacts represent upper level course work, where students should be demonstrating mastery of this outcome, a mere 2.4% of students demonstrated exemplary performances.

## Outcome 2: Empathy (UD Core)



A similar trend is seen for outcome 2. For “Empathy,” only 2 artifacts of 126 scored “Exemplary” on the rubric. This represents a mere 1.59% of students. Thus, we see that students at the upper level are not demonstrating mastery of this outcome. The overwhelming majority are performing at the “Beginning” or “Emerging” levels at a time in their college careers when they should be prepared to matriculate and thus demonstrate mastery of the learning outcomes we claim our curriculum delivers.

### Outcome 3: Openness (UD Core)



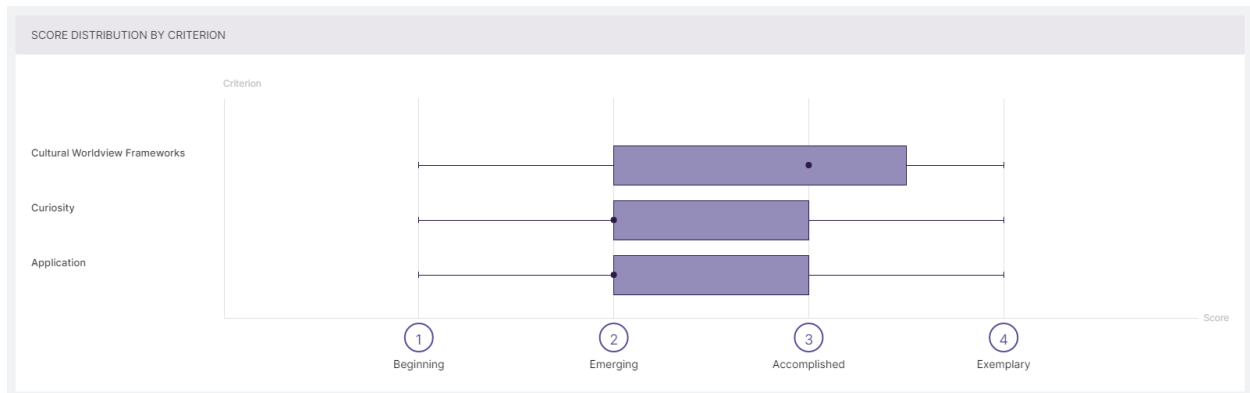
Only 4 of 126 or 3.17% of artifacts scored at the “Exemplary” level, repeating the poor performances seen in the previous two outcomes.

In general, the data reflects trends to be expected. The median score across all outcomes overall was 2, and the median score increased across all outcomes, except for “Cultural Self-Awareness.” Yet the frequency data is clear: very few artifacts demonstrated “Exemplary” performances, which is suggestive of a failure of our students to demonstrate mastery of any of the outcomes under Goal A.

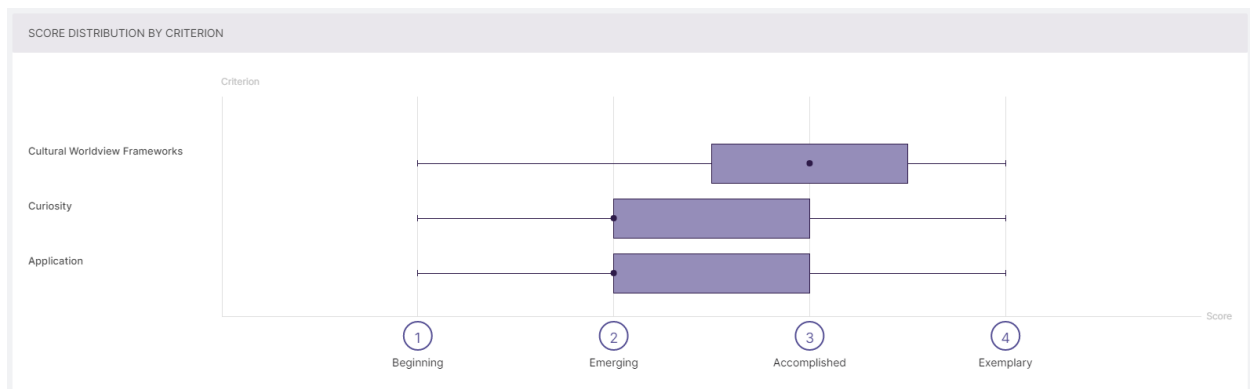
## Goal B (Other)

This is by far the largest goal of the Diversity competency. Even so, 48.64% of artifacts were scored over the three-day evaluation period.

### Overall Score Distribution by Outcome



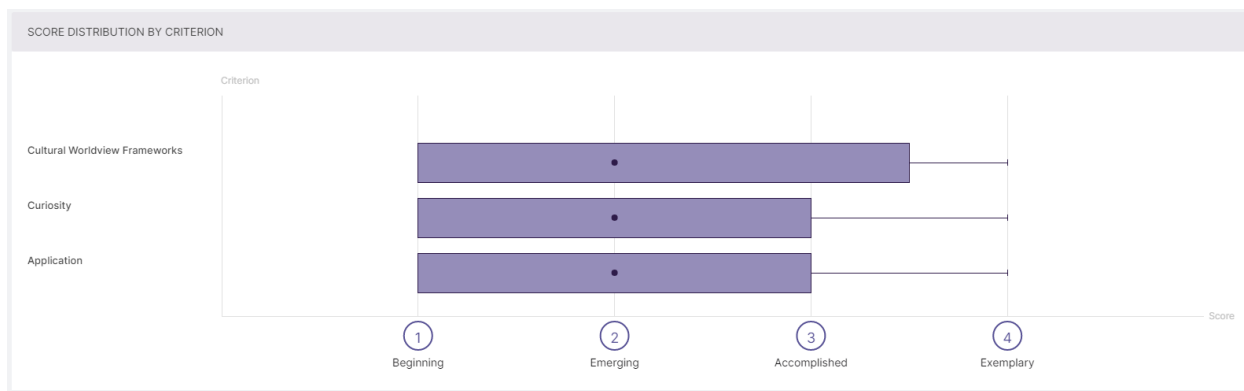
### Lower Division Overall by Outcome



	Max Score	Min Score	Median Score	# of submissions
Cultural Worldview Frameworks	4	1	3	505
Curiosity	4	1	2	505
Application	4	1	2	505

This goal has much greater representation across the lower division in terms of course selection, so the data are more reliable.

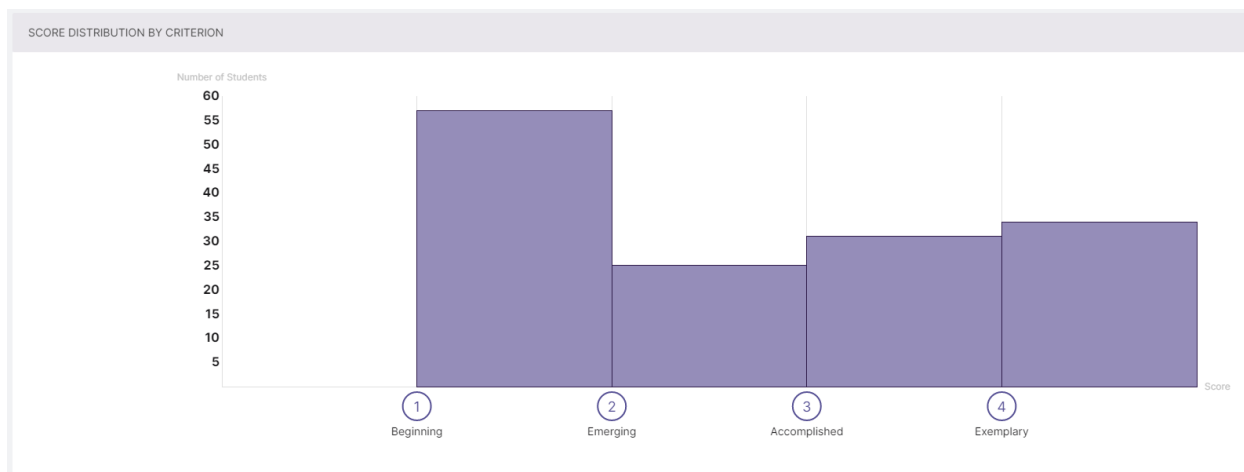
### Upper Division by Outcome



	Max Score	Min Score	Median Score	# of submissions
Cultural Worldview Frameworks	4	1	2	147
Curiosity	4	1	2	147
Application	4	1	2	147

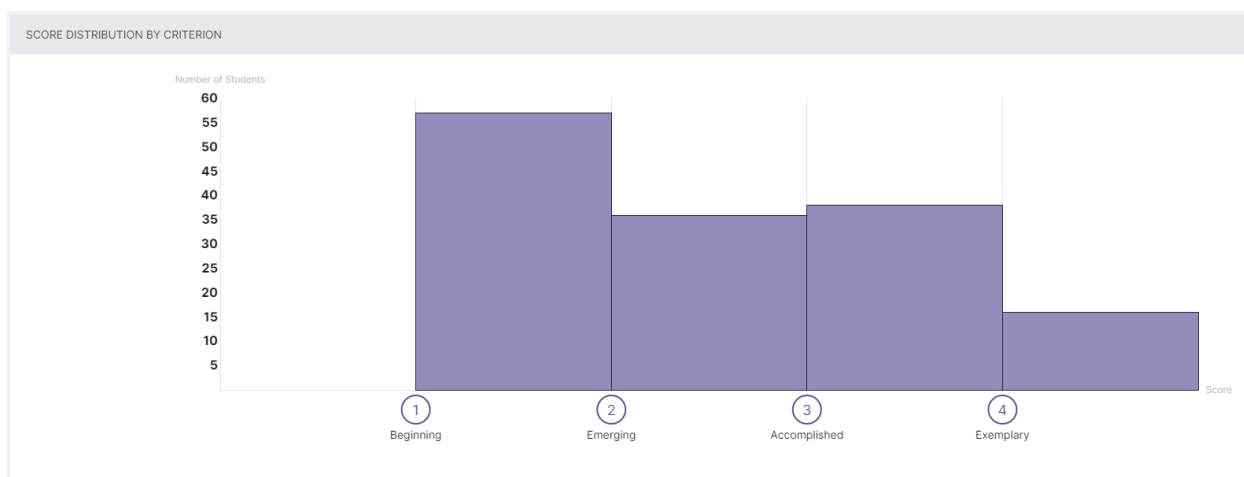
Overall, the upper division scores for Diversity Goal B are stagnant in comparison with the LD Core courses, with the only notable exception being a drop in median score in the Cultural Worldview Frameworks, from a 3 in the LD Core to a 2 in the UD Core. One would expect, or hope, that at the upper division, students would be performing much higher on the rubric, with a significant proportion of students scoring at the 3 (Accomplished) or 4 (Exemplary) levels.

## Outcome 1: Cultural Worldview Frameworks (UD Core)



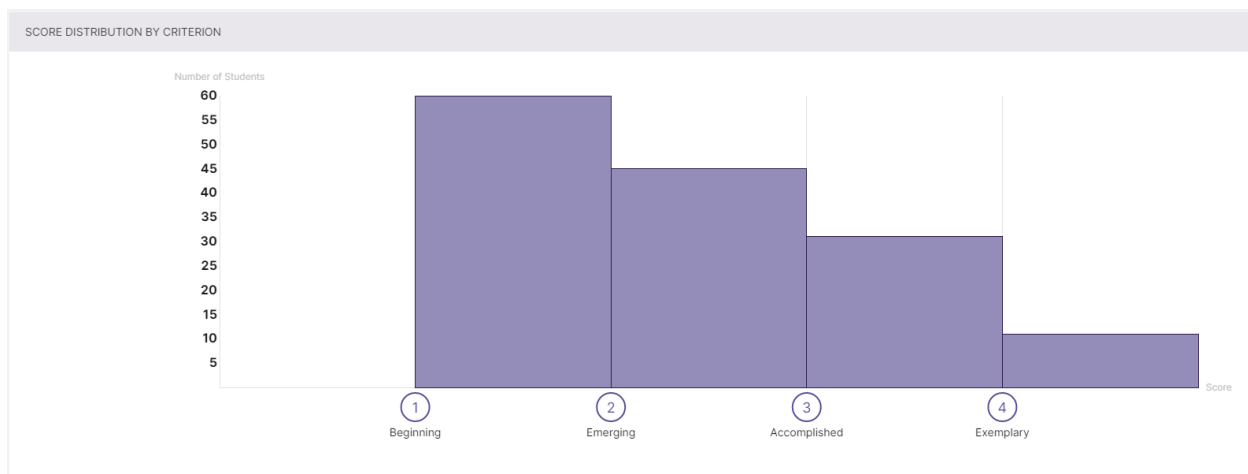
Of note, 34 of a total 147 artifacts scored registered at level 4, or exemplary. 23.13% of students at the upper level scored “Exemplary.” Taking into account the “Accomplished” level, we see that 44.28% of artifacts were scored at a 3 or higher, indicating that nearly half of those artifacts scored in upper division courses demonstrated a level of mastery we’d expect to see at this level. That is reassuring. We can do better, but we are clearly on the right track in this area.

## Outcome 2: Curiosity (UD Core)



Outcome 2 does not fare as well. Here only 10.88% scored “Exemplary.” Including the “Accomplished” level, we see 36.73% scoring at a 3 or higher. This is a much better record than we see with goal A, but there is definite room for improvement. The majority of students are performing at the “Beginning” or “Emerging” levels at time in their college careers when they should be prepared to matriculate and thus demonstrate mastery of the learning outcomes.

### Outcome 3: Application (UD Core)



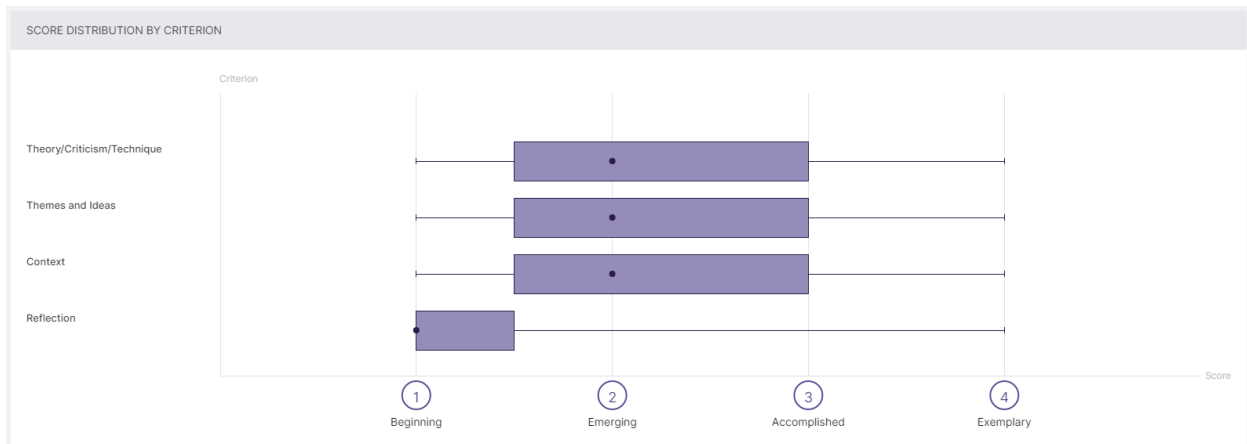
Only 11 of 147 or 7.48% of artifacts scored at the “Exemplary” level, repeating the poor performances seen in goal A. This outcome, specifically, had many issues when it came to applying the rubric to the artifacts. Faculty scorers noted that many times the student work provided simply did not address the outcome and so they could not apply the rubric to the student work. In addition, the distinction between level 3 and 4 was unclear. In general, the opinion of the score team was that this rubric was poorly designed. This poor design lead to difficulties in scoring and so the data reflects these issues.

In general, the data reflects trends to be expected. Students do progress, modestly, from the lower to upper division. Although under the Cultural World Frameworks outcome students do well, for the other two outcomes, especially “Application,” very few artifacts demonstrated “Exemplary” performances, which is suggestive of a failure of our students to demonstrate mastery.

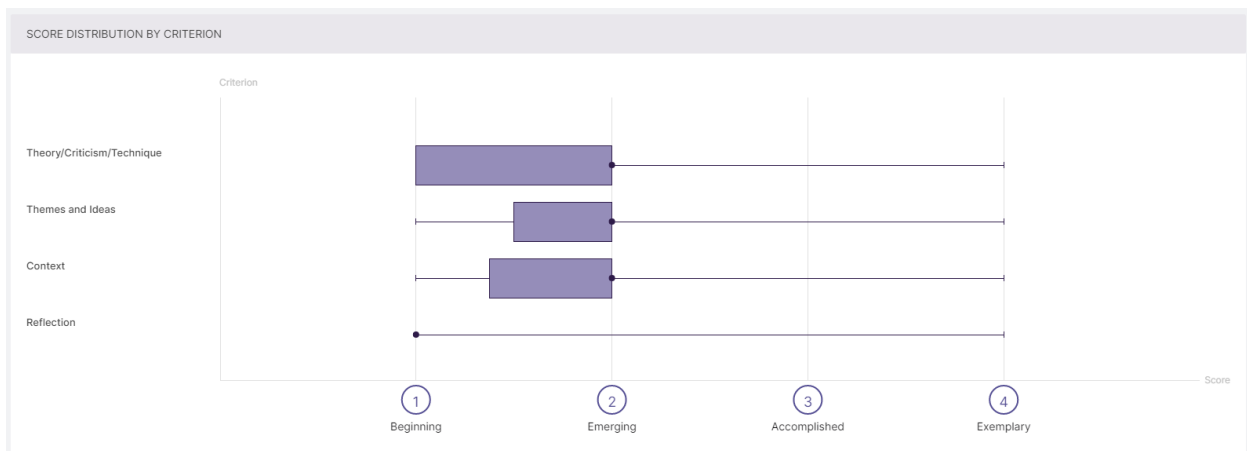
## Goal C (Creative Works)

51.50% of artifacts were scored over the three-day evaluation period.

### Overall Score Distribution by Outcome



### Lower Division Overall by Outcome

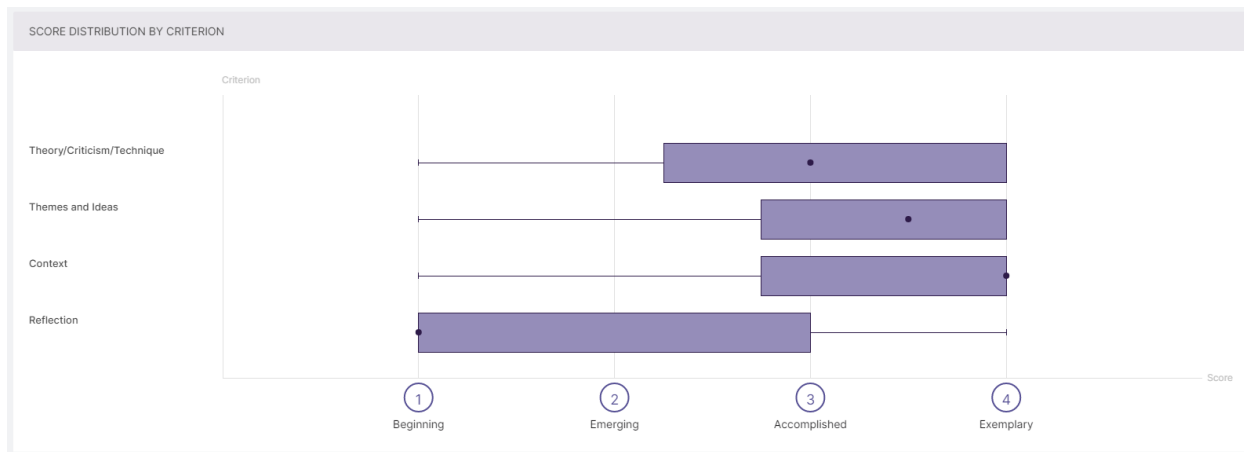


	Max Score	Min Score	Median Score	# of submissions
Theory/Criticism/Technique	4	1	2	386
Themes and Ideas	4	1	2	368
Context	4	1	2	368

Reflection	4	1	1	368
------------	---	---	---	-----

This goal, also, has much greater representation across the lower division in terms of course selection, so the data are more reliable.

### Upper Division by Outcome

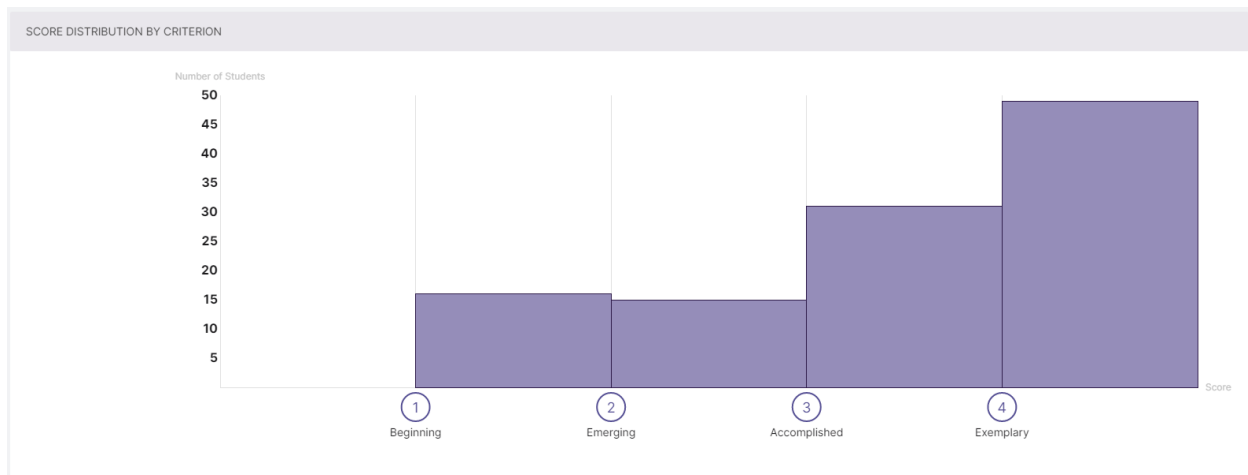


### UPDATE

	Max Score	Min Score	Median Score	# of submissions
Theory/Criticism/Technique	4	1	3	111
Themes and Ideas	4	1	4	111
Context	4	1	4	111
Reflection	4	1	1	111

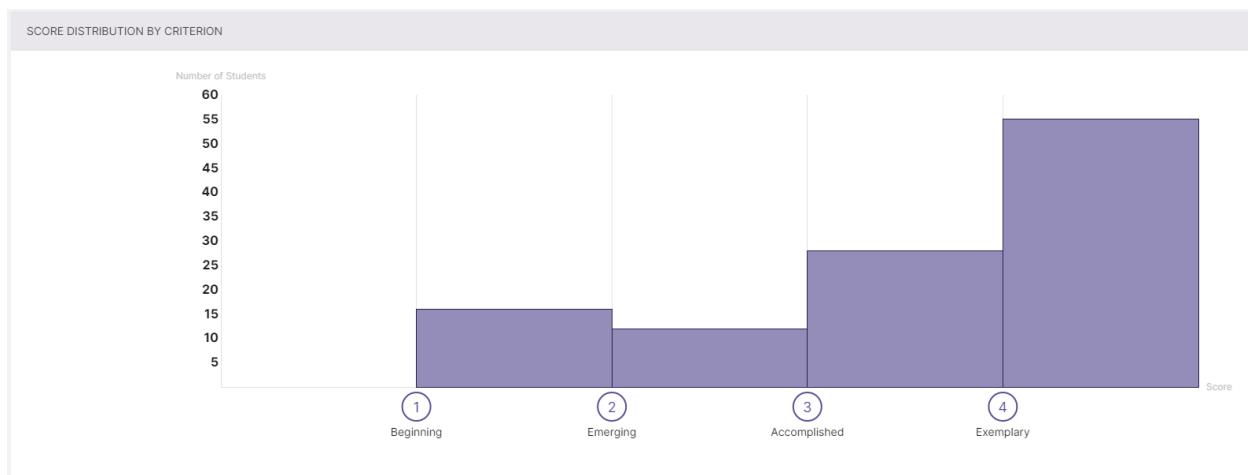
Overall, the upper division scores for Diversity Goal C show significant growth. Note the median increases to 3, 4, and 4 for Theory/Criticism/Technique, Themes and Ideas, and Context, respectively. The jump to a median score of 4 in the latter two outcomes is impressive. Yet there is notable stagnation for the “Reflection” outcome.

## Outcome 1: Theory/Criticism/Technique (UD Core)



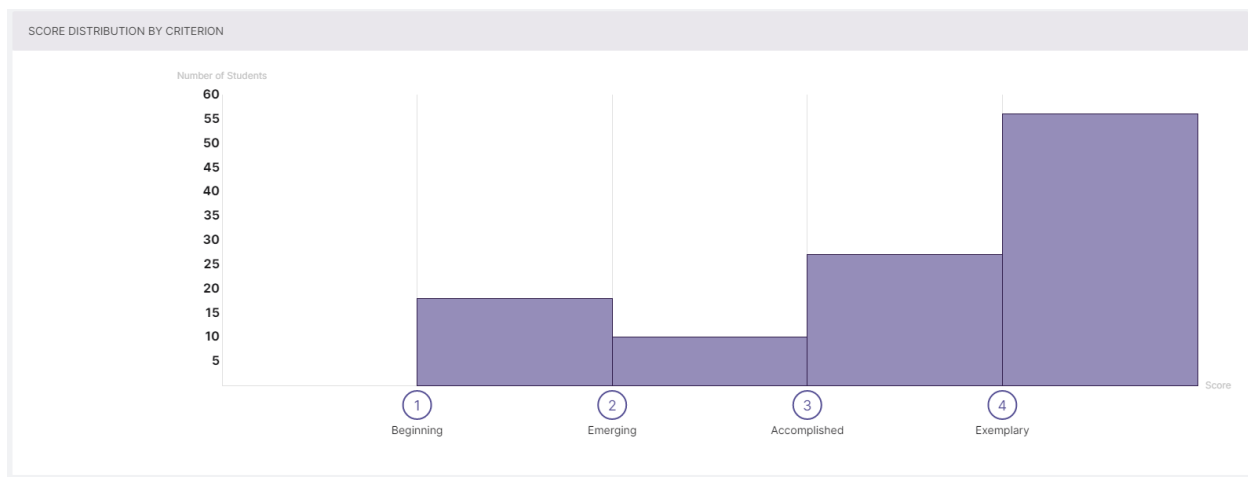
Of note, 49 of a total 111 artifacts scored registered at level 4, or exemplary. 44.14% of students at the upper level scored “Exemplary.” Taking into account the “Accomplished” level, we see that 72.97% of artifacts were scored at a 3 or higher, indicating that well over half of those artifacts scored in upper division courses demonstrated an level of mastery we’d expect to see at this level. This is indeed reassuring.

## Outcome 2: Themes and Ideas (UD Core)



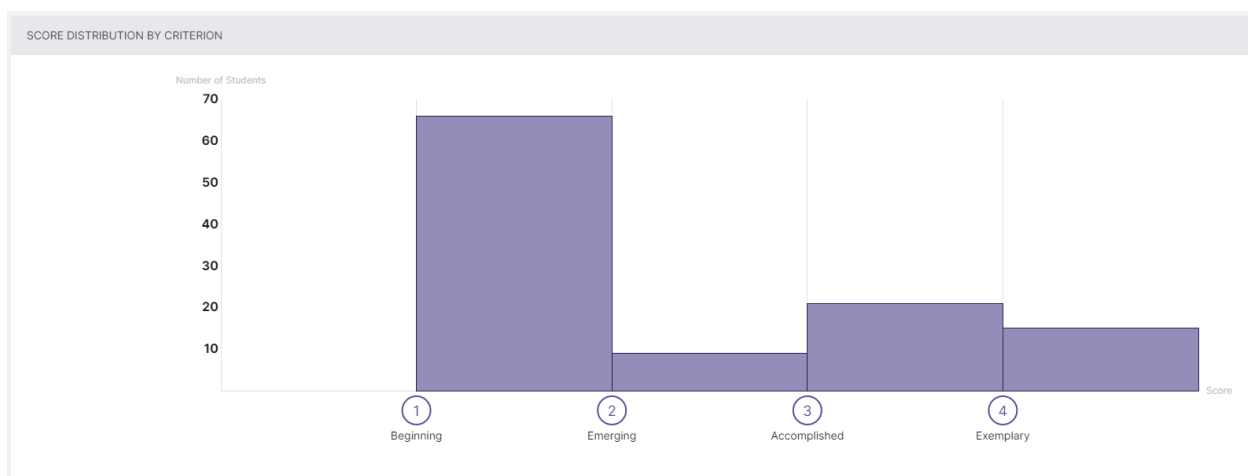
Outcome 2 fares as well as outcome 1. Here 55 of 111, or 49.55% scored “Exemplary.” Including the “Accomplished” level, we see 74.77% scoring at a 3 or higher. This is a much better record than we see with goals A or B. The majority of students are demonstrating mastery.

### Outcome 3: Context (UD Core)



Outcome 3 fares as well. Here 56 of 111, or 50.45% scored “Exemplary.” Including the “Accomplished” level, we see 75.68% scoring at a 3 or higher.

### Outcome 4: Reflection (UD Core)



Outcome 4 is an outlier. In fact, it is the inverse of the previous 3 outcomes. Here 66 of 111, or 59.46% scored “Beginning” with only 15 of 111 or 13.51% scoring “Exemplary.”

In general, the data reflects trends to be expected, with one notable outlier. Students do progress, in some regards markedly, from the lower to upper division. The frequency data is unambiguous. Although under the Reflection outcome we see a problematic result, stagnation at the “Beginning” level.

## **V. Conclusions and Recommendations**

The data support the following observations and recommendations:

- 1) Faculty participation improved. Efforts by the Office of Assessment need to remain focused on this issue so this trend of improved participation continues.
- 2) Given that poorly chosen or designed assignments pose a problem in terms of generating representative data, a problem frequently noted by the score teams (See Appendix B), pre-cycle training needs to focus on assignment design and needs to be readily accessible and more widely used by faculty.
- 3) With respect to student learning: growth was noted in some areas. These numbers afford us the opportunity to set benchmarks against which to evaluate the data from the second cycle of assessment of the Diversity competency. However, the fact that less than half of students at the upper division in Goals A and B scored “accomplished” or higher, and that often less than a quarter or fifth of upper level students scored “Exemplary” is an issue and indicates that there is room for improvement. Students at the upper division should demonstrate a greater amount of mastery at higher rates if our curriculum is to be adjudged impactful.
- 4) Goal A showed troubling results with less than 5.00% of students at the upper division scoring “Exemplary” on the rubric. This is problematic.
- 5) Goal C demonstrated an exemplary trend among students. At the upper level, for all outcomes except “Reflection” almost half of all students scored “Exemplary.”
- 6) Diversity outcomes tend to assess for “dispositions” and so seek expressions of such “dispositions,” rather than proof of knowledge. Outcomes like “empathy,” “openness,” and “curiosity,” are difficult, if not impossible, to demonstrate unambiguously. Thus, aside from other issues pertaining to inculcating a disposition of thinking in our students, assessing this ability is problematic. The nature of these outcomes and the flaws in the extant assignments mean these data are not highly reliable.
- 7) Given scorer feedback, the rubrics need to be revisited. Issues to be addressed include the ability of non-experts to apply the rubric in the assessment exercises as carried out at UCA and more precise and consistent wording in the rubrics to better standardize expectations among scorers.

## Appendix A: ICC calculations via SPSS

### Goal A

#### Intraclass Correlation Coefficient

	Intraclass Correlation <sup>b</sup>	95% Confidence Interval		F Test with True Value 0			
		Lower Bound	Upper Bound	Value	df1	df2	Sig
Single Measures	.602 <sup>a</sup>	.520	.672	4.011	272	272	.000
Average Measures	.751 <sup>c</sup>	.684	.804	4.011	272	272	.000

Two-way mixed effects model where people effects are random and measures effects are fixed.

- The estimator is the same, whether the interaction effect is present or not.
- Type A intraclass correlation coefficients using an absolute agreement definition.
- This estimate is computed assuming the interaction effect is absent, because it is not estimable otherwise.

### Goal B

#### Intraclass Correlation Coefficient

	Intraclass Correlation <sup>b</sup>	95% Confidence Interval		F Test with True Value 0			
		Lower Bound	Upper Bound	Value	df1	df2	Sig
Single Measures	.167 <sup>a</sup>	.066	.265	1.402	362	362	.001
Average Measures	.286 <sup>c</sup>	.123	.419	1.402	362	362	.001

Two-way mixed effects model where people effects are random and measures effects are fixed.

- The estimator is the same, whether the interaction effect is present or not.
- Type A intraclass correlation coefficients using an absolute agreement definition.
- This estimate is computed assuming the interaction effect is absent, because it is not estimable otherwise.

### Goal C

#### Intraclass Correlation Coefficient

	Intraclass Correlation <sup>b</sup>	95% Confidence Interval		F Test with True Value 0			
		Lower Bound	Upper Bound	Value	df1	df2	Sig
Single Measures	.687 <sup>a</sup>	.632	.735	5.479	431	431	.000
Average Measures	.814 <sup>c</sup>	.774	.847	5.479	431	431	.000

Two-way mixed effects model where people effects are random and measures effects are fixed.

- The estimator is the same, whether the interaction effect is present or not.
- Type A intraclass correlation coefficients using an absolute agreement definition.
- This estimate is computed assuming the interaction effect is absent, because it is not estimable otherwise.

## Appendix B: Scorer Comments in AQUA

### Goal A

The assignment asks students to compare views of past and current Greek society, define concepts, and describe a few theories. The last question offers possibilities for discussion of diversity to demonstrate self-awareness and perhaps empathy, but openness isn't addressed.

The assignment doesn't ask students to discuss their own cultural beliefs, imagine one's self as another, or address their openness to interacting with those of other cultures, so while these ideas are sometimes mentioned, there's no expectation to address them.

The assignment doesn't ask students to discuss their own cultural beliefs, imagine one's self as another, or address their openness to interacting with those of other cultures, so while these ideas are sometimes mentioned, there's no expectation to address them.

The assignment doesn't ask students to discuss their own cultural beliefs, imagine one's self as another, or address their openness to interacting with those of other cultures although Question #4 does prompt some students to address these questions.

The assignments have one question that pertains to cultural diversity, which is generally answered in a paragraph. Five or six sentences doesn't allow scores to get into the 3 or 4 range.

Using the one question that addresses aspects of cultural diversity is typically answered in one paragraph. This makes it difficult to score in the 3-4 range.

Using the one question that addresses aspects of cultural diversity is typically answered in one paragraph. This makes it difficult to score in the 3-4 range.

The assignment has two questions asking students if they agree with what Mia says about poor people, and if our country is doing a good job at elevating voices of the unheard and if the student is doing so. These questions are generally answered in 2-4 sentences, which does not allow for scores of 3 or 4. Reaching a 2 is difficult as well since students are not asked to address the items on the rubric.

Many of the essays score low because students often take diversity to mean representing the different types of student groups on campus--those in a certain major or have a certain interest. Some do discuss LGBTQ+ or religious signs and mostly all discuss the language of the signs but only to the extent of identifying them as such and whether or not they belong to these groups. They usually don't discuss this diversity in terms of empathy or openness though.

While this assignment is for a 1000 level course, the assignment itself forces students to think more critically/analytically about the topic. As such, higher scores have been selected.

The essay prompt includes "Do they serve to encourage diversity beyond what is demanded from the law, or to limit it?" but was not addressed by the student, and while there was a discussion concerning "are some groups advantaged or disadvantaged by it?" the student never defined the groups they referred to.

This assignment asks students to discuss groups that may be advantaged or disadvantaged by certain laws, but that doesn't necessarily lead to the points of the diversity rubric. Like this particular student discusses a cross-burning and a band name that deals with disadvantaged groups but the student isn't discussing these cases in relation to the rubric.

While the assignment calls for students to rate themselves on a scale of how well they understand the language and cultural diversity of their clients, the reflection usually does not discuss their knowledge of how their experience has shaped their cultural awareness (rubric item #1), or their level of empathy with clients (item #2). Item #3, Openness, is inherent in the scale but there's usually no discussion of this aspect such that they can meet scores above 1. I have a feeling that if most of these students were directed to specifically discuss the points of the rubric as it relates to cultural competence checklist, they would be scoring in the 3 and 4 range easily.

These artifacts are obviously from a presentation, but I don't have the presentation--just the powerpoints, which are visual aids and notes. Difficult to apply the rubric to notes.

## Goal B

The instructor's assignment is fine but the student's response does not seem to have anything to do with diversity of culture.

This assignment does not seem to have much to do with cultural diversity.

The student's work has effectively no application of Core Diversity Goal B.

This assignment does not seem to have anything to do with diversity of culture.

This submission has very little to do with cultural diversity.

I have not the foggiest idea what this has to do with Diversity Core B.

This assignment does not appear to have much (if any) relationship with the UCA Core B objectives. I am not sure if this is due to the nature of the assignment, or if the student just chose not to consider other cultures.

This assignment is N/A due to it being in French.

This assignment does not appear to have much (if any) relationship with the UCA Core B objectives.

I see nothing whatsoever about different cultures in this student's response.

This submission does not appear to have much (if any) relationship with the UCA Core B objectives. Since there are no assignment instructions, I do not know if this is due to the nature of the assignment or to the performance of the student.

This is a scattershot response without much depth.

Without instructions, it is hard to be certain, but this assignment should be N/A. It does not appear to have much (if any) relationship with the UCA Core B objectives.

While this response is admittedly not singularly domestic, it is not multi-cultural in that it focuses rather exclusively upon Japan.

I do not understand what is happening here.

This assignment does not correlate Diversity Rubric B

Mi Espanol es no Bueno.

## Goal C

I am guessing that the assignment does not ask for a reflection.

This response was not accompanied by a prompt from the student's professor, which makes it a bit more challenging to know what was asked of the student.

This assignment is not written to be a good fit for the assignment.

This assignment does not ask for the specific categories on rubric C

This assignment does not really ask for a reflection

This assignment does not address the rubric.

The challenge is that the rubric refers to ""the work(s), but the student's response is not to a work but their own example that could be related to a work. The only mention of a work is the title of this apparent student blog thread. But this SLO reads ""[r]ecognizes the general context(s) [of the work(s)] but cannot connect context to its impact..."". Given that the student is writing about her own town (and not the town of the source subject), it is unclear to me that there is enough there there."

This assignment does not really address Rubric C.

This item does not address any features of Rubric C

This assignment is not an ideal fit for Rubric C.

This assignment requires brief descriptions of many different poems -- it's not a great fit for earning higher scores on this rubric

I'm not sure this assignment asks for a personal reflection.

This assignment does not include the notes supporting this Powerpoint.

The prompt is for a presentation, but the artifact is the slides only. For this reason, it is very difficult to assign this example from an upper-division diversity-designated course. A presentation would normally be delivered along with the slides, filling in many of the gaps that the slides in isolation do not address. I think these types of artifacts are problematic unless we also have a recording of the spoken presentation."

The title of this presentation includes the word ""colored,"" which surprised me. I discussed with my assessment partner whether the student(s) were unaware of the baggage associated with the term-- or for that matter how the term was invoked during the presentation accompanying the slideshow. One of the major challenges with assessing this artifact is that the accompanying assignment referred to a presentation and gave the presenters leeway to add additional information to the slideshow. Yet since as an assessor I do not get to see the presentation, I am really not able to make sense of the links between the images and textual fragments in the same way that the professor (or students attending the presentation) would have. Minus a video recording of the presentation, I do not see this type of artifact as particularly useful for diversity assessment."

The presentation slideshow is not accompanied by the delivered presentation. For that reason, it is not the ideal artifact for assessing with the diversity rubric since it does not address the gaps left out by the absence of the spoken presentation that was included in the original rhetorical situation.

One other problem with many of these slideshow artifacts is that most of them do not analyze or even focus on another work (or works). Since both SLOs 1 and 3 specifically refer to work(s) as part of their framing for assessment, these slideshows are generally outside the scope of what the rubrics are designed to assess.

Being generous, therefore, only SLOs 2 and 4 potentially relate to most of these slideshow artifacts."

Because this assignment is for a presentation, the submitted slideshow does not provide the comprehensive overview of the rhetorical situation in which it was used. For that reason, any assessor will miss the clarifications that would have been there when the slideshow was displayed along with the presentation. For that reason, this type of artifact does not work well for this rubric."

This assignment requires very brief descriptions of 6 different features -- it's not a great fit for earning higher scores on this rubric

This assignment, again, seems to require very superficial responses to many pieces. It is unlikely that students will be able to earn scores higher than 1, because they aren't discussing any piece in any depth (nor comparing them).

The assignment seems to ask for very superficial responses to many pieces.

Assignment not provided.

This item is missing pages.

I would like to see the assignment for this item. It is not clear that the assignment is a good match for the rubric.

This response was not accompanied by the assignment sample from the professor.

On SLO 1, the response incorrectly designates specific national cultures as high-context and low-context cultures. For this reason, I assigned the response a 3 for knowing the concepts but opted not to assign it a 4 for inaccurately applying the concepts to specific national cultures.