

Assessment of Critical Thinking at the University of Arizona

Fall 2024

Overview

The assessment of Critical Thinking (CT) at the University of Arizona (U of A) is the third of four ABOR-requested assessments of our general education program. The objective of this rubric-based assessment was twofold: 1. to measure student achievement of CT, and 2. to identify areas for improvement in student learning to strengthen our general education program in the area of critical thinking.

The U of A faculty defined Critical Thinking in the following way:

Interdisciplinary Definition of Critical Thinking

Critical thinking includes the three skills of argumentation, evaluation of evidence, and meaning making along with the disposition of reflective thinking.

Critical thinking can be divided into two categories: skills and dispositions.

Skills: synthesize, evaluate, create, innovate, justify, interpret, integrate, demonstrate, argumentation, analysis of assumptions and biases, evaluation of evidence, question formulation, sound reasoning/evidence-based reasoning, problem-solving, comprehensive exploration of ideas/concepts, meaning-making, application, practice, adaptation

Dispositions: Open-mindedness, reflective (metacognitive) thinking - analysis of one's own assumptions and biases, perspective-taking, self-regulated judgment, agency/self-confidence, curiosity, prudence (careful consideration)

Achievement of this outcome is best measured through institutional coursework that incorporates CT into the curriculum. The U of A sample included 238 artifacts from 4 200-level General Education courses, and 263 CT-identified artifacts from the UNIV 301 student-created ePortfolios for a total of 501 artifacts.

Who Participated in Scoring Student Work?

- Over 40 faculty, staff, and doctoral students participated at different stages of the assessment which included focus groups and a year-long working group.
- 21 participants were trained as evaluators to score students' work. They represented many of the colleges and disciplinary areas across campus.

What Student Samples Were Assessed?

- Student work was collected from two sources: 4 different lower-division, general education courses whose signature assignments were identified by the instructor as meeting the criteria for Critical Thinking, and UNIV 301, our GE capstone course in which students identify artifacts from GE courses that align with the CT outcome.

- From a collection of over 800 artifacts, 501 were selected for evaluation to be sure that we had a representative sample across classes.

How Was the Rubric Created?

A collective of our faculty created the U of A's definition of Critical Thinking that was shared with the Tri-University working group. Representatives from our faculty worked with the Tri-University group to build a common rubric for the 3 institutions to use in this assessment. See Appendix A.

How was Student Work Assessed?

- To ensure reliability, all reviewers were calibrated on the rubric prior to the scoring process. Each artifact was evaluated twice, with a third reading taking place if the first two scores showed a difference greater than 1.
- When reviewing the artifacts, if evidence of a particular criterion was not present, the scorer could mark n/a. In these cases, the artifact was not included in the average score for that criterion, explaining why the *n* for each criterion is different.

What Did We Learn?

The table below shows the frequency of student scores for each of the rubric criteria. It is clear from these data that most of the students are meeting or exceeding expectations in all areas of Critical Thinking. Two areas, Evaluation of Evidence and Reflective Thinking, appear to be the weakest in performance. As Reflective Thinking is a new emphasis in our GE curriculum, we look forward to that area improving the next time we assess this outcome. These findings are aligned or slightly better than those at [Oklahoma State University, 2017](#), [Texas A & M, 2021](#), [Eastern Illinois University 2023](#) and [Hawaii-Pacific University, 2019](#).

	Exceeds	Meets Expectations	Developing
Argumentation (n=495)	89%		
	33%	56%	11%
Evaluation of Evidence (n=464)	77%		
	27%	51%	22%
Analysis/ Interpretation/ Reflection (n=495)	82%		
	32%	50%	18%
Conclusions/ Outcomes/ Meaning-Making (n=499)	88%		
	28%	59%	13%

When we disaggregated the data by course vs ePortfolio artifact, student artifacts from the classes scored a little higher in all areas than the ePortfolio artifacts. This difference was significant in the Meaning-Making criterion.

Because our general education curriculum is only in its second year, the ePortfolio process is new, and only a small percentage of students have yet to take it. It is meant for students to reflect back on their learning in their GE curriculum and highlight assignments that aligned with the various GE outcomes. Over time, we expect to see an increase in performance in this process mainly due to more experienced teaching practices throughout the entire GE program. Some initiatives that we should consider moving forward include:

- Offering workshops on writing effective signature assignments in critical thinking, including how to scaffold an assignment.
- Developing Faculty Learning Communities based on best practices in teaching critical thinking.
- Consider other workshops or forums that can help faculty with the teaching of critical thinking in a variety of areas.