



Critical Thinking Artifact Assessment Report 2023-2024

Executive Summary

- In AY 2023-24, 90 Critical Thinking artifacts were scored by two separate raters using a revised version of the AAC&U Critical Thinking VALUE rubric. The rubric consists of four criteria:
 - Explanation of Issues
 - Evidence
 - Student's Position
 - Conclusions and Related Outcomes
- Each criterion was rated on a five-point scale: Capstone (4), Milestones (3 and 2), Benchmark (1), and Not Present (0). It also included the option to mark an artifact as unscorable.
- Conclusions and Related Outcomes yielded the highest mean score for differences in rating by criterion at $\bar{x} = 1.02$. Evidence and Student's Position yielded the lowest standard deviation at $sd = 0.85$. When averaging the total differences in ratings, all criteria yielded a mean score of $\bar{x} = 0.96$ and a standard deviation of $sd = 0.87$.
- Overall, almost half of the artifacts were scored on average at Milestone (3) (43.7%), followed by a comparable portion of the artifacts being at Milestone (2) (39.1%). Capstone (4) ratings were representative of 8.0% of artifacts, Benchmark (1) ratings were representative of 9.2% of artifacts, and no artifacts resulted in a Not Present (0) rating.
- For overall normalized ratings by criteria, 2.21 to 1.94 was the mean score range, with an overall mean score of 2.04 on the five-point scale from 4.00 to 0.00.
 - The criterion with the highest mean score was Explanation of Issues ($\bar{x} = 2.21$); 8.0% of Critical Thinking artifacts were rated at the Capstone (4) level, which was the highest number of Capstone (4) ratings among all criteria.
 - The criterion with the second highest mean score was Evidence ($\bar{x} = 2.08$).
 - The criteria with the lowest mean scores were Conclusions and Related Outcomes ($\bar{x} = 2.02$) and Student's Position ($\bar{x} = 1.94$).

Introduction

Critical Thinking was assessed during the 2023-24 academic year (AY) using a modified version of the Association of American Colleges and Universities (AAC&U) Valid Assessment of Learning in Undergraduate Education (VALUE) rubric (see Appendix A). The AAC&U VALUE rubrics were developed by teams of educational professionals and include the most frequently identified criteria of learning for different learning outcomes. Washburn University (Washburn) implements performance assessments using a modified version (2015) of the AAC&U VALUE rubrics for assessing Critical Thinking every three years. Artifacts in written format are collected from students in EN 300: Advanced College Writing and are scored by two or more independent raters using the Outcomes Assessment Projects (formerly known as Aqua) by Watermark software platform.

Review Process

Washburn faculty were invited to attend the calibration training conducted in-person on May 16, 2024. The 12 faculty who attended the training were assigned 90 artifacts collected from Fall 2023 and Spring 2024 EN 300: Advanced College Writing courses. The artifacts were assigned to be reviewed by two independent raters on four criteria: Explanation of Issues, Evidence, Student's Position, and Conclusions and Related Outcomes. These four criteria were scored on the five-point scale of Capstone (4), Milestones (3 and 2), and Benchmark (1), with the additional Not Present (0) for scoring criteria that did not meet Benchmark (1) level performance. Reviewers could also assign the status of unscorable to those artifacts that were not appropriate for the purpose of critical thinking assessment.

Of the 12 reviewers who participated in the artifact review process, 11 scored 15 artifacts and one scored 9 artifacts. Three artifacts were designated as unscorable, which resulted in a total of 87 artifacts reviewed two times each, for a total of 348 scores across 174 reviews.

Results

Differences by Criterion

The AAC&U Critical Thinking VALUE rubric defines Critical Thinking as “a habit of mind characterized by the comprehensive exploration of issues, ideas, artifacts, and events before accepting or formulating an opinion or conclusion.” The rubric contains four criteria: Explanation of Issues, Evidence, Student's Position (perspective, thesis/hypothesis), and Conclusions and Related Outcomes (implications and consequences). The written artifacts were rated on these four criteria on a five-point scale consisting of Capstone (4), Milestones (3 and 2), Benchmark (1), and Not Present (0).

The 87 Critical Thinking artifacts were rated by two reviewers on four dimensions for a total of 348 scores. These scores were reviewed for each of the four criteria to examine differences in ratings per criterion. Table 1, on the following page, provides the descriptive statistics for the differences in ratings by criterion.

Table 1. Descriptive Statistics for Differences in Ratings by Criterion

	Explanation of Issues	Evidence	Student's Position	Conclusions and Related Outcomes	Total Difference
Mean	0.95	0.90	0.95	1.02	0.96
St. Dev.	0.90	0.85	0.85	0.90	0.87
Min	0.00	0.00	0.00	0.00	0.00
Max	3.00	3.00	3.00	3.00	3.00

The mean difference was greatest for Conclusions and Related Outcomes with an average difference of rating at $\bar{x} = 1.02$. The mean difference was smallest for Evidence with an average difference in rating at $\bar{x} = 0.90$. The standard deviation was smallest – indicating the results are clustered more closely around the mean and thus more dependable – for Evidence and Student’s Position at $sd = 0.85$. The minimum rating was 0.00 (or no difference) across all criteria, whereas the maximum rating was 3.00 across all criteria. The total mean difference across all aggregated criteria was $\bar{x} = 0.96$, and the standard deviation was $sd = 0.87$.

Distribution of Scores

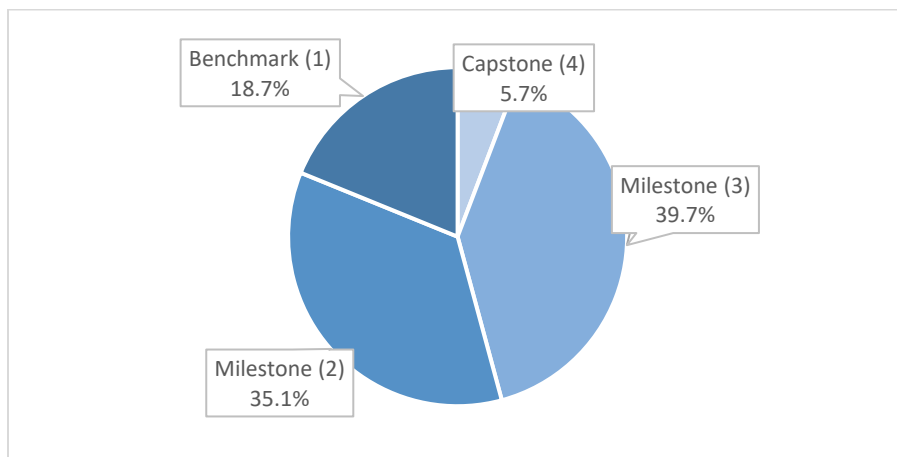
The scores from the two raters were averaged to provide a normalized score for each Critical Thinking artifact. Table 2, below, provides the overall averaged rating of 348 scores for 87 Critical Thinking artifacts. The ranges of average scores were defined as 4.00-3.01 Capstone, 3.00-2.01 and 2.00-1.01 Milestones, and 1.00-0.01 Benchmark. The criteria that were not rated are designated Not Present.

Table 2. Descriptive Data and Statistics for Overall Averaged Ratings

	Capstone (4) 4.00 - 3.01	Milestone (3) 3.00 - 2.01	Milestone (2) 2.00 - 1.01	Benchmark (1) 1.00 - 0.01	Not Present 0.00	Mean (sd)
Overall (n = 348)	20 (5.7%)	138 (39.7%)	122 (35.1%)	65 (18.7%)	3 (0.9%)	2.04 (0.81)

Most ratings (39.7%) were scored at Milestone (3), followed by Milestone (2) (35.1%). Twenty (2) ratings (5.7%) were at Capstone (4), 65 (18.7%) were at Benchmark (1), and three (0.9%) were at Not Present (0). See Figure 1 below for a visual representation (excluding Not Present).

Figure 1. Distribution of Scores for Overall Critical Thinking Artifacts



Ratings by Criterion

The distribution of average ratings and descriptive statistics for each of the four criteria are in Table 3, below. The table shows how the artifacts ($n = 87$) were rated on each of the four criteria (total of 348 scores).

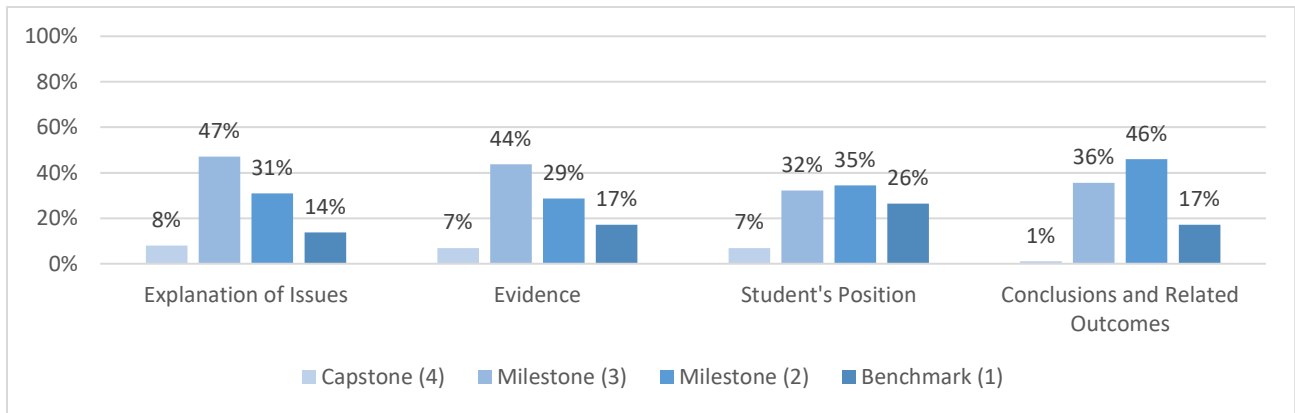
Table 3. Descriptive Data and Statistics for Overall Averaged Ratings by Criterion

	Capstone (4)	Milestone (3)	Milestone (2)	Benchmark (1)	Not Present (0)	Mean (<i>sd</i>)
	4.00 - 3.01	3.00 - 2.01	2.00 - 1.01	1.00 - 0.01	0.00	
Explanation of Issues	7 (35.0%)	41 (29.7%)	27 (22.1%)	12 (18.4%)	0 (0%)	2.21 (1.02)
Evidence	6 (30.0%)	38 (27.5%)	25 (20.5%)	15 (23.1%)	3 (100%)	2.08 (1.08)
Student's Position	6 (30.0%)	28 (20.3%)	30 (24.6%)	23 (35.4%)	0 (0%)	1.94 (1.05)
Conclusions and Related Outcomes	1 (5.0%)	31 (22.5%)	40 (32.8%)	15 (23.1%)	0 (0%)	1.95 (0.97)
Overall ($n = 348$)	20 (100%)	138 (100%)	122 (100%)	65 (100%)	3 (100%)	2.04 (1.03)

For Explanation of Issues and Evidence, the ratings fell within the 2.21 to 2.08 mean range, indicating that these artifacts' scores were on the lower end of the Milestone (3) range, on average. Although the highest percent of scores for both Explanation of Issues and Evidence fell in the Milestone (3) range (47.1% and 43.7%, respectively), both also had a disproportionately low number of scores in the Capstone (4) range (8.0% and 6.9%, respectively), which likely contributed to mean scores on the lower end of the Milestone (3) range. It is worth noting that all criteria saw a disproportionately low number of scores in the Capstone (4) range for AY 2024. Although Student's Position and Conclusions and Related Outcomes fell within the 1.95 to 1.94 mean range, both criteria fell on the higher end of the Milestone (2) range. Most scores for both Student's Position and Conclusions and Related Outcomes fell within the Milestone (2) range (34.5% and 46.0%, respectively). Finally, Student's Position accounted for most of the scores (35.4%) which fell within the Benchmark (1) range.

See the bar chart on the following page for a visual representation of the distribution of ratings by each criterion (excluding Not Present). Student's Position was the only criterion with a normal distribution of ratings. Explanation of Issues and Evidence were positively skewed while Conclusions and Related Outcomes was negatively skewed. See Figure 2 on the following page for a visual representation.

Figure 2. Distribution of Scores by Criterion



Areas of Consideration and Limitations

The Critical Thinking artifacts for 2023-24 were not rated by a third party due to an error during the administration of third-party ratings; the results from 2023-24 cannot be compared across years.

Reasonable efforts were made to collect Critical Thinking artifacts from students enrolled in EN 300: Advanced College Writing, a university requirement for Junior level students during the Fall 2023 and Spring 2024 terms. A random sampling was not used to select artifacts for review; however, the submissions were voluntary. Given that this course is a requirement for all Junior level students, general assumptions could be made about the proficiency level of all students at Washburn in Critical Thinking.

Many of the 12 faculty reviewers were new to the artifact scoring process, whereas in past years, the same faculty were asked to participate in scoring. This may have resulted in more variation in scores than in prior years.

CRITICAL THINKING VALUE RUBRIC

for more information, please contact value@aacu.org



The VALUE rubrics were developed by teams of faculty experts representing colleges and universities across the United States through a process that examined many existing campus rubrics and related documents for each learning outcome and incorporated additional feedback from faculty. The rubrics articulate fundamental criteria for each learning outcome, with performance descriptors demonstrating progressively more sophisticated levels of attainment. The rubrics are intended for institutional-level use in evaluating and discussing student learning, not for grading. The core expectations articulated in all 15 of the VALUE rubrics can and should be translated into the language of individual campuses, disciplines, and even courses. The utility of the VALUE rubrics is to position learning at all undergraduate levels within a basic framework of expectations such that evidence of learning can be shared nationally through a common dialog and understanding of student success.

Definition

Critical thinking is a habit of mind characterized by the comprehensive exploration of issues, ideas, artifacts, and events before accepting or formulating an opinion or conclusion.

Framing Language

This rubric is designed to be transdisciplinary, reflecting the recognition that success in all disciplines requires habits of inquiry and analysis that share common attributes. Further, research suggests that successful critical thinkers from all disciplines increasingly need to be able to apply those habits in various and changing situations encountered in all walks of life.

This rubric is designed for use with many different types of assignments and the suggestions here are not an exhaustive list of possibilities. Critical thinking can be demonstrated in assignments that require students to complete analyses of text, data, or issues. Assignments that cut across presentation mode might be especially useful in some fields. If insight into the process components of critical thinking (e.g., how information sources were evaluated regardless of whether they were included in the product) is important, assignments focused on student reflection might be especially illuminating.

Glossary

The definitions that follow were developed to clarify terms and concepts used in this rubric only.

- **Ambiguity:** Information that may be interpreted in more than one way.
- **Assumptions:** Ideas, conditions, or beliefs (often implicit or unstated) that are "taken for granted or accepted as true without proof." (quoted from www.dictionary.reference.com/browse/assumptions)
- **Context:** The historical, ethical, political, cultural, environmental, or circumstantial settings or conditions that influence and complicate the consideration of any issues, ideas, artifacts, and events.
- **Literal meaning:** Interpretation of information exactly as stated. For example, "she was green with envy" would be interpreted to mean that her skin was green.
- **Metaphor:** Information that is (intended to be) interpreted in a non-literal way. For example, "she was green with envy" is intended to convey an intensity of emotion, not a skin color.

CRITICAL THINKING VALUE RUBRIC

*for more information, please contact value@aacu.org
Revised 2015 for use at Washburn University USLO Assessment*

Definition

Critical thinking is a habit of mind characterized by the comprehensive exploration of issues, ideas, artifacts, and events before accepting or formulating an opinion or conclusion.

Evaluators are encouraged to assign a zero to any work sample or collection of work that does not meet benchmark (cell one) level performance.

	Capstone 4	Milestones		Benchmark 1	Not Present 0
		3	2		
Explanation of issues	Issue/problem to be considered critically is stated clearly and described comprehensively, delivering all relevant information necessary for full understanding.	Issue/problem to be considered critically is stated, described, and clarified so that understanding is not seriously impeded by omissions.	Issue/problem to be considered critically is stated but description leaves some terms undefined, ambiguities unexplored, boundaries undetermined, and/or backgrounds unknown.	Issue/problem to be considered critically is stated without clarification or description.	Evaluators are encouraged to assign a zero to any work sample or collection of work that does not meet benchmark (cell one) level performance.
Evidence <i>Selecting and using information to investigate a point of view or conclusion</i>	Information is taken from source(s) with enough interpretation/evaluation to develop a comprehensive analysis or synthesis.	Information is taken from source(s) with enough interpretation/evaluation to develop a coherent analysis or synthesis.	Information is taken from source(s) with some interpretation/evaluation, but not enough to develop a coherent analysis or synthesis.	Information is taken from source(s) without any interpretation /evaluation.	Evaluators are encouraged to assign a zero to any work sample or collection of work that does not meet benchmark (cell one) level performance.
Student's position (perspective, thesis/hypothesis)	Specific position (perspective, thesis/hypothesis) is sophisticated, taking into account the complexities of an issue. Limits of position (perspective, thesis/hypothesis) are acknowledged. Others' points of view are synthesized within position (perspective, thesis/hypothesis).	Specific position (perspective, thesis/hypothesis) takes into account the complexities of an issue. Others' points of view are acknowledged within position (perspective, thesis/hypothesis).	Specific position (perspective, thesis/hypothesis) acknowledges different sides of an issue.	Specific position (perspective, thesis/hypothesis) is stated, but is simplistic and obvious.	Evaluators are encouraged to assign a zero to any work sample or collection of work that does not meet benchmark (cell one) level performance.
Conclusions and related outcomes (implications and consequences)	Conclusions and related outcomes (consequences and implications) are logical and reflect student's informed evaluation and ability to place evidence and perspectives discussed in priority order.	Conclusions are logically tied to a range of information, including opposing viewpoints; related outcomes (consequences and implications) are identified clearly.	Conclusions are logically tied to information (because information is chosen to fit the desired conclusion); some related outcomes (consequences and implications) are identified clearly.	Conclusions are inconsistently tied to some of the information discussed; related outcomes (consequences and implications) are oversimplified.	Evaluators are encouraged to assign a zero to any work sample or collection of work that does not meet benchmark (cell one) level performance.