Assessment Brief DRAFT
Critical Thinking—AAC&U Values Rubric Findings
Prepared for AOAC, Fall 2010

Part I: Executive Summary

A. **General Education Outcome**
    Students will demonstrate foundational knowledge of the important ideas and methods of different ways of knowing. The focus of this brief will be on the critical thinking component of general education.

B. **Background**
    The purpose is to seek information to share with faculty and administrators to enhance opportunities for student learning experiences. AOAC members found the American Association of Colleges and Universities (AAC&U) new "Values Rubrics" to be useful tools for assessing sample work.

    In Spring 2010, over 140 student artifacts were collected from 12 different courses with permission from faculty and students. Thirty of the artifacts were randomly selected from the 140 submitted artifacts. Of the artifacts randomly selected, 4 were from 100-level courses, 9 were from 200-level courses, 10 were from 300-level courses, and 5 were from 400-level courses.

C. **Major Findings**
    *
    *   Intraclass correlation coefficient and Cronbach's Alpha levels for the rubric for critical thinking were .62 and .89 respectively.

D. **Conclusions**
    *
    *   A respectable level of interrater reliability and rater consistency were achieved, and in general, the rubric appears to be a sound method of assessment.

Part II: Assessment Brief

A. **Introduction**
   1. **Background**
   One of the Objectives for General Education at Millersville University is critical thinking competence. In this component, students are expected to "analyze and interpret," "investigate, evaluate, and apply," and "develop the necessary tools of critical thinking, inquiry, and diplomacy."

   In Spring 2010, the Faculty Senate Academic Outcomes Assessment Committee (AOAC) designed a process to assess general education competencies (also known as University Outcomes) as they aligned with general education objectives. This new process evolved after a review of the prior general education assessment strategies resulted in the need to enhance triangulation of evidence and to utilize more direct evidence of student learning such as student work from courses. The purpose is to seek information to share with faculty and administrators to enhance opportunities for student learning experiences. AOAC members found the American Association of Colleges and Universities (AAC&U) new "Values Rubrics" to be useful tools for assessing sample work.

   2. **Problem Statement**
   AOAC members want to:
      i.   explore the feasibility of the AAC&U Values Rubrics, specifically, critical thinking, and how it informs the respective student learning outcome; and
      ii.   develop and better understanding of how well Millersville students demonstrate critical thinking in various course levels in their educational experience.

   3. **Justification**

   AOAC with support of Planning, Assessment and Analysis has access to indirect evidence of students' perceptions of their development of critical thinking (NSSE) and nationally-normed critical and analytical thinking instruments (CLA and CAAP Critical Reasoning). Little evidence has been collected of actual student work to inform our understanding of critical thinking. In addition, the nationally-normed instruments are expensive and difficult to gain reasonable student response rates. The use of direct evidence collected from purposive sampling of courses provides an opportunity to explore student learning experiences at Millersville and triangulate that information with other measures of the outcome.

B. **Information Source**
   In order to test the Critical Thinking rubric, written papers were collected from sections of Biology, Business, Chemistry, Mathematics, Music, Psychology, and First Year Seminar classes in the spring semester of 2010. Papers of students who signed the consent forms were coded (n = 140).

C. Major Findings

1. Data Summary

Intraclass Correlation Coefficient and Cronbach's Alpha levels for the rubric were tested on each of the five constructs for the written papers. The number of observed scores and the mean score are also included in the table below for each construct. The second table includes a comparison of means scores among construct and course level as identified by the course number from which the work was done. The third shows the changes of means scores for each increase of course level, examining if the constructs change over students' progression through their program.

Table 1: Critical Thinking Mean Scores and Reliability by Construct

	Number of Scores	Mean Score	Intraclass Correlation	Cronbach's Alpha
Explanation of Issues	150	2.4	.57	.87
Selecting and Using Evidence	135	2.1	.72	.93
Influence of Context	150	2.0	.60	.88
Student's Position	150	2.2	.57	.87
Conclusions and Related Outcomes	150	2.1	.61	.89
Overall	**735**	**2.2**	**.62**	**.89**

Table 2: Critical Thinking Mean Scores by Course Level and Construct

	Course Level			
	100	**200**	**300**	**400**
Explanation of Issues	1.3	2.6	2.7	2.6
Selecting and Using Evidence	0.6	2.3	2.4	2.5
Influence of Context	0.8	2.4	2.2	2.1
Student's Position	1.1	2.5	2.5	1.9
Conclusions and Related Outcomes	**0.8**	**2.3**	**2.4**	**2.1**
Overall	**0.9**	**2.4**	**2.4**	**2.2**

Table 3: Changes in Critical Thinking Mean Scores with Increasing Course Level
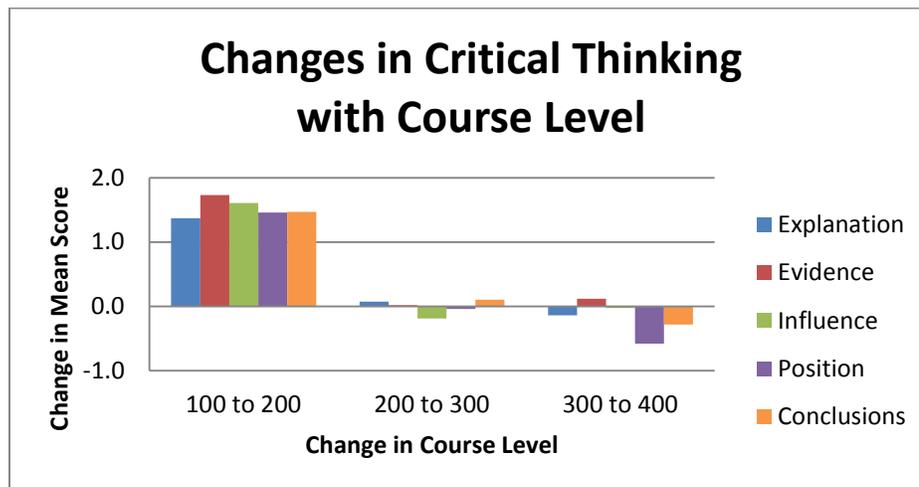
Note that only one course was classified as a 100 level course

2.   Discussion

- A Cronbach's Alpha reliability coefficient of .80 or higher is considered "good" in most social science research situations; this value is "used to rate the internal consistency (homogeneity) or the correlation of the items in a test." The intraclass correlation coefficient is "used to measure inter-rater reliability for two or more raters," and a value greater than .60 is considered "acceptable" (a value just greater than .50 is sometimes considered "barely acceptable").   In using and testing this rubric, the five raters were able to achieve respectable levels of interrater reliability, and the rubric appears to be sound.

- "Explanation of Issues" and "Student's Perspective" were weaker items in terms of reliability and consistency of raters. It might be helpful to revise these two sections of the rubric, providing greater clarity of what is being assessed and greater distinction between the levels of scoring.

- It was not the aim of this study to examine the current levels of critical thinking of students. However, if they were to be analyzed in the future, it is recommended that more than one course be used in each class level, since the largely picture-based assignment in the only 100 level course provided an unfairly low average score for that classification. It might also be noted that "Selecting and Using Evidence" was the only construct to always increase as course level progressed.

- Three student artifacts were not rated on the construct "Selecting and Using Evidence," and these were each from a MATH 310 course which did not require the use of sources. If this rubric will be use for assessment in the classroom, it may

have to be adjusted for certain departments, courses, or projects depending on the scope of the work.

- The mean scores which quantify the level of students' critical thinking skills would be most helpful if there were goals and clearer meaning to the numeric result. If the rubric will be used for actual assessment in courses, it is recommended that the numerical scores be associated with levels of competence that have relevant meaning.

3. **Conclusions**

- If this rubric is intended for usage in the university, it is suggested that certain constructs be clarified and that stronger connections be made between numeric scores and course-specific meaning.

- Once the final version of this rubric is complete, it should be available to interested faculty to be used for critical thinking assessment beyond the courses examined here, and faculty should be trained how to use the rubric.
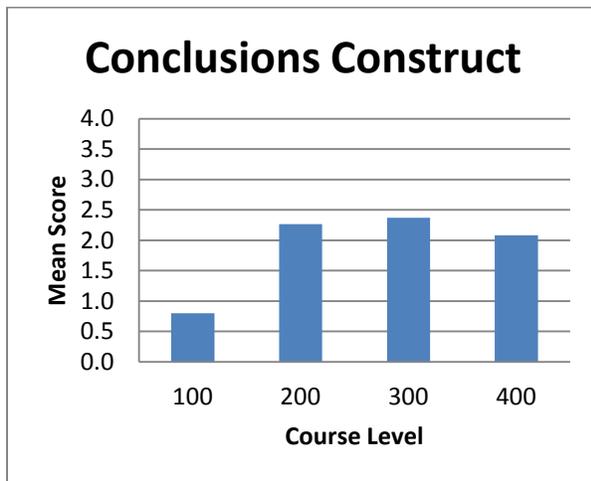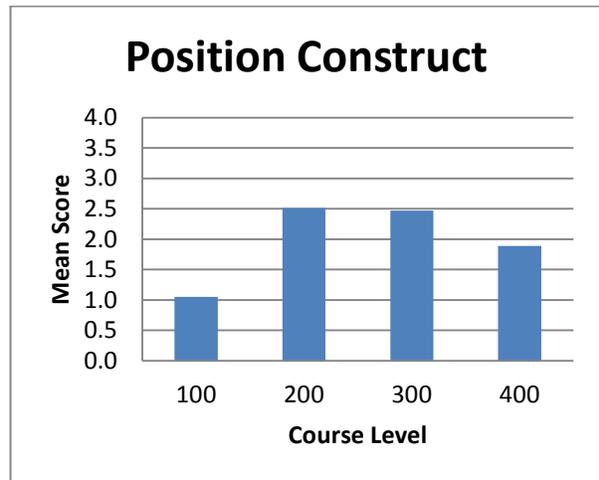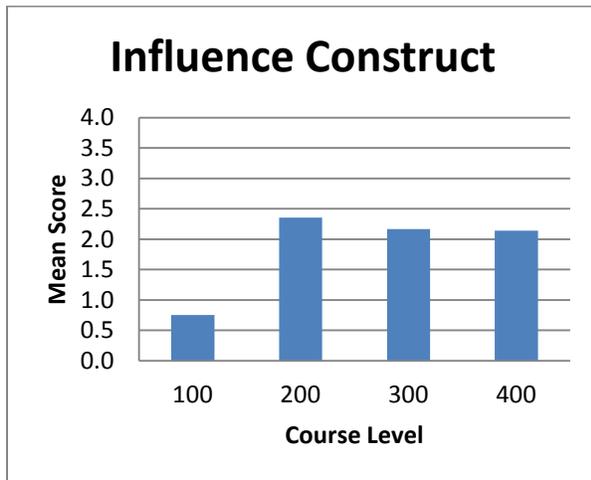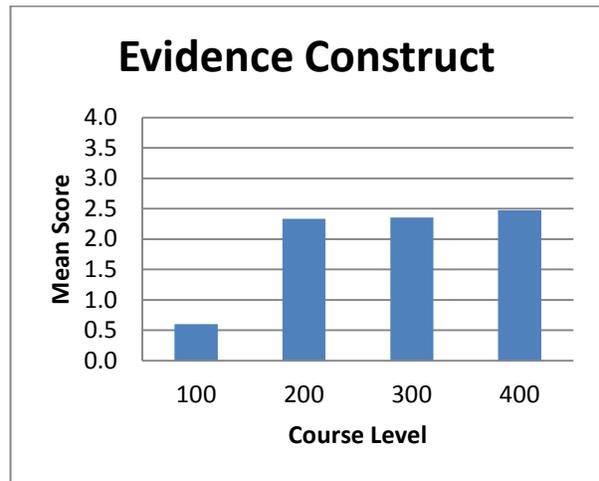
Part III: Appendices

CRITICAL THINKING RUBRIC FOR ASSESSMENT OF STUDENT WORK (AAC&U VALUES RUBRIC)

	Capstone 4	Milestones 3	Milestones 2	Benchmark 1
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.
Evidence *Selecting and using information to investigate a point of view or conclusion*	Information is taken from source(s) with enough interpretation/evaluation to develop a comprehensive analysis or synthesis. Viewpoints of experts are questioned thoroughly.	Information is taken from source(s) with enough interpretation/evaluation to develop a coherent analysis or synthesis. Viewpoints of experts are subject to questioning.	Information is taken from source(s) with some interpretation/evaluation, but not enough to develop a coherent analysis or synthesis. Viewpoints of experts are taken as mostly fact, with little questioning.	Information is taken from source(s) without any interpretation/evaluation. Viewpoints of experts are taken as fact, without question.
Influence of context and assumptions	Thoroughly (systematically and methodically) analyzes own and others' assumptions and carefully evaluates the relevance of contexts when presenting a position.	Identifies own and others' assumptions and several relevant contexts when presenting a position.	Questions some assumptions. Identifies several relevant contexts when presenting a position. May be more aware of others' assumptions than one's own (or vice versa).	Shows an emerging awareness of present assumptions (sometimes labels assertions as assumptions). Begins to identify some contexts when presenting a position.
Student's position (perspective, thesis/hypothesis)	Specific position (perspective, thesis/hypothesis) is imaginative, 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.
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.	Conclusion is logically tied to a range of information, including opposing viewpoints; related outcomes (consequences and implications) are identified clearly.	Conclusion is logically tied to information (because information is chosen to fit the desired conclusion); some related outcomes (consequences and implications) are identified clearly.	Conclusion is inconsistently tied to some of the information discussed; related outcomes (consequences and implications) are oversimplified.

Figures 1-5: Scores for Critical Thinking Constructs by Course Level

Explanation Construct

Mean Score vs Course Level (100, 200, 300, 400)

Evidence Construct

Mean Score vs Course Level (100, 200, 300, 400)

Influence Construct

Mean Score vs Course Level (100, 200, 300, 400)

Position Construct

Mean Score vs Course Level (100, 200, 300, 400)

Conclusions Construct

Mean Score vs Course Level (100, 200, 300, 400)
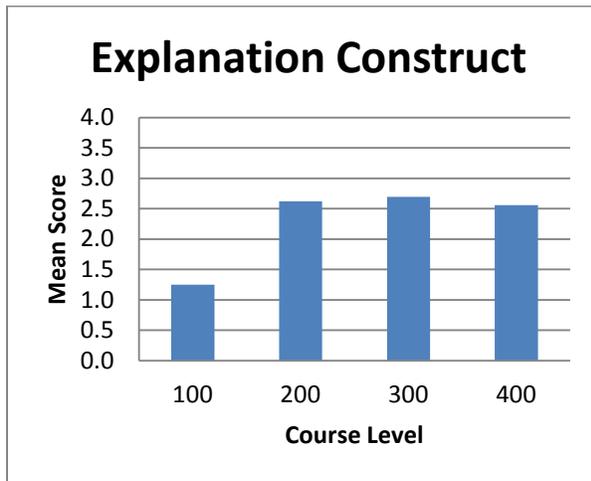
Note that only one course was classified as a 100 level course

Table 4: Critical Thinking Score Statistics by Construct

	Explanation of Issues	Evidence	Influence of Context	Student's Position	Conclusion and Outcomes	Overall
Number of Scores	150	135	150	150	150	**735**
Mean Score	2.4	2.1	2.0	2.2	2.1	**2.2**
Standard Error of Mean	0.07	0.08	0.07	0.07	0.07	**0.03**
Standard Deviation	0.80	0.88	0.83	0.79	0.88	**0.85**
Intra-Class Correlation	0.57	0.72	0.60	0.57	0.61	**0.62**
Cronbach's Alpha	0.87	0.93	0.88	0.87	0.89	**0.89**

Glossary of Statistical Terms

Mean – the average score

Variance – a measurement of how much (far) the scores vary around the mean score

Standard Deviation – the square root of the variance
   - Measures the same thing as variance: how far scores are from the mean score

Standard Error of the Mean – a measurement of how much the group mean scores vary around the
          total mean score (the standard deviation of the group means)

Interrater Reliability/Agreement

- A measure of consistency and usefulness of the rubric
- The extent to which independent raters agree on a rubric score and to which rubric scores are consistent across raters
- In this assessment, measured by Intraclass Correlation Coefficient and Cronbach's Alpha

Intraclass Correlation Coefficient (ICC)

- A measure of interrater reliability that describes how strongly scores from the same rater resemble each other. It is a value from 0 (no rater reliability) to 1 (complete rater reliability).
- Mathematically, it's the proportion of the total variance that's due to variability between raters
- Can also be interpreted as a measure of between group differences or within group similarity

Cronbach's Alpha

- A specific type of ICC that is test of internal consistency (ranges on same scale as ICC)
- Measures how well a set of raters measure a single, latent (covert) construct
- Estimates how strongly the score obtained from the actual panel of the raters correlates with the score that would have been obtained from another random sample of raters