

## ACADEMIC ASSESSMENT RECORD 2006-2007

YOUR COLLEGE/SCHOOL: CLAS

DEGREE PROGRAM COVERED BY THIS PLAN: BA in Measurement

% OF YOUR FACULTY MEMBERS WHO CONTRIBUTED TO THE DEVELOPMENT OF THIS PLAN:  
100%

NAME OF YOUR DEPARTMENT'S ASSESSMENT REPRESENTATIVE: Keston Fulcher

TODAY'S DATE (FOR TRACKING): 9/12/2007 (that's approximately when I would turn in the Record corresponding to the 2006-2007 plan)

Congruency of Outcomes to Guiding Documents:

Our department established intended learning outcomes for students that are congruent with elements of CNU's Vision 2010, and our own departmental goals.

Regarding the Strategic Plan our intended learning outcomes load on at least two of its goals: IA: Provide an intellectually challenging and dynamic liberal learning curriculum and IB: Support learning that cultivates critical and innovative thinking. Our criteria for success for all our learning outcomes are intended to be challenging, setting the bar high for our students and faculty. Our 2<sup>nd</sup> and 3<sup>rd</sup> learning outcomes relate to students creating and evaluating research studies. Both of these activities require critical thinking skills.

The goals for our department are (1) Provide students with foundational knowledge in measurement theory, (2) Train students to apply their knowledge of measurement by creating and evaluating research, and (3) Prepare students for jobs in measurement or graduate school. Our first learning outcome is directly related to our 1<sup>st</sup> departmental goal; we're assessing foundational knowledge in measurement. Our 2<sup>nd</sup> and 3<sup>rd</sup> learning outcomes directly relate to the 2<sup>nd</sup> departmental goal – students' ability to create and evaluate research. Finally, all of our intended learning outcomes are related to a student's future prospects. The foundation we emphasize and the ability to create and evaluate research are major components of any job or graduate school curriculum in measurement. In essence, assessment results of our intended learning outcomes help indicate this department's success relative to CNU's Strategic Plan and our own departmental goals.

**Comment [khf1]:** Item 1 on the Review form. To score a four, explicitly make the connection between the intended learning outcomes, Vision 2010 and one other guiding document.

**I. Intended Learning Outcome 1:** Students will identify and distinguish among basic measurement concepts and practices including reliability, validity, performance assessment, sampling, and fairness in testing.

**Comment [khf2]:** Item 3 on review form. Note that the outcome is aligned with what a student should know, think, or do.

A. First Means of Assessment/Criteria for Success: We used a pre-post design using ETS' Concepts in Measurement Test to assess how much majors have gained on this outcome from their early junior to senior years. As juniors, students took the test at the beginning of METH 321, Intro to Research Methods. As seniors, they took the test at the end of METH442, Senior Seminar in Research and Statistics. The test is a requirement for both courses and participation is 100% (85 out of 85 juniors and 70 out of 70 seniors). However, because only 62 students matched at both times, the other students were dropped from the analysis. The test is administered per ETS' instructions. All students appeared to give good effort, but four juniors and three seniors did not have time to complete all items. The test was chosen by the department because (1) it covers all of the topics stated in this learning outcome, (2) the publisher provides considerable validity evidence for its use with college students, and (3) it has national norms. Criteria for success: We would like our students' average post score overall and each component therein to be 30 percentage points higher than the pre score (e.g., from the 30<sup>th</sup> percentile to the 60<sup>th</sup> percentile), which would be a slight improvement over past gains for most areas.

**Comment [khf3]:** Note that all of the 1<sup>st</sup> means of assessment throughout the Assessment Record are direct measures such as tests or performance assessments (e.g., faculty ratings of students' reports, papers, oral presentations, portfolios, etc). Indirect measures such as surveys are fine for a second means of assessment.

**Comment [khf4]:** Item 7 on review form: Note thorough explanation of data collection.

**Comment [khf5]:** Item 6 on review form. Note the rationale given for selecting the test.

**Comment [khf6]:** Item 8 on review form: note that not only are criteria for success stated but a rationale is given.

*Results:*

Gains on ETS' Concepts in Measurement Test

	Overall	Reliability	Performance assessment	Validity	Sampling	Fairness in testing
Juniors (Fall 2004, n =62)	27%ile	35%ile	20%ile	29%ile	30%ile	22%ile
Seniors (Spring 2006, n = 62)	57%ile	75%ile	31%ile	60%ile	55%ile	44%ile
<b>Current (2006) Difference in %iles</b>	<b>+ 30</b>	<b>+ 40</b>	<b>+ 11</b>	<b>+ 31</b>	<b>+ 25</b>	<b>+ 22</b>
Last Year (2005) Difference in %iles	+ 24	+ 30	+ 13	+ 35	+19	+ 24
Two Yrs Ago (2004) Difference in %iles	+ 26	+ 37	+ 9	+ 28	+28	+ 30

**Comment [khf7]:** Item 9 on review sheet. Note that results are clearly displayed. Providing multiple years of results provides useful longitudinal information. Tables may not be necessary if results are very simple.

B. Second Means of Assessment/Criteria for Success: All seniors (n = 70) took a departmental graduate exit survey. This survey was co-designed by department faculty and the Director of Assessment & Evaluation. A survey item corresponds to each of the topics stated in the intended learning outcome: reliability, performance assessment, validity, sampling, and fairness in testing. Students are asked to rank order which of these areas they learned the most in the major. Because the department values learning in all of these topics, we would prefer that no topic is

**Comment [khf8]:** Review item 5: to score a 4, the dept. must include a 2<sup>nd</sup> means of assessment for three outcomes.

ranked consistently lower than others, which we define as .5 or more difference in average ranking.

### Results

Rank Order of Measurement Concepts

Area	Avg. Rank
Reliability	2.6
Sampling	2.8
Fairness in testing	2.9
Validity	2.9
Performance assessment	3.8*

#### Interpretation and Use of Results:

The findings were mixed. We met our criteria for success overall (an improvement of 30 %ile points); however, students did not improve to that degree across all areas. As in years past, reliability was a strength with students improving 40 %ile points. Students just met our expectations in the area of validity. On the other hand, students improved merely 11 %ile points in performance assessment. Students' gains in sampling and fairness in testing were also below expectations but to a much lesser degree. The pattern of strengths and weaknesses has not shifted dramatically over the past years. Clearly, students have consistently made smaller gains in performance assessment than in any other area.

**Comment [khf9]:** Item 10 on review form: The conclusions drawn correspond directly with the stated learning outcome.

Regarding the second means of assessment, students average rankings varied more than the department would like. Most salient, students ranked performance assessment much lower than the other areas (about a full point). This finding lends convergent evidence that students are not learning as much about performance assessment.

The last three year's results suggest that the department should focus on enriching the curriculum related to performance assessment. We have attempted to do just this. At the beginning of the 2005 fall semester, the department analyzed the curriculum in relation to performance assessment. We found that less teaching time and fewer assignments were devoted to the topic. In response we have implemented four strategies to improve this area.

**Comment [khf10]:** Item 12 on review form: The numerous references in this Assessment Record to past assessments indicate a strong commitment to assessment.

1. Awareness: Students are told at the beginning of METH 321 that there will be increased emphasis on performance assessment.
2. An extra three class meetings in METH 321 are devoted to teaching performance assessment. To make room, sessions on non-parametric statistics were reduced from six to three. Although a difficult decision, faculty felt justified increasing class time on performance assessment at the expense of non-parametric statistics.
3. More items about performance assessment are added to the mid-term exam, in essence weighting the topic more heavily.
4. Faculty in the department are encouraged to reinforce the topic of performance assessment in their classes.

**Comment [khf11]:** Item 11 on review form. This is a good example of a department using results to improve their curriculum. Note how a weakness was identified and that the interventions directly target that area.

Because these changes affected mostly juniors this year, these interventions likely did not have any bearing on this year's senior results. Next year, however, the seniors will have received the interventions and the department anticipates that their scores on performance assessment will improve substantially.

**II. Intended Learning outcome 2:** Seniors will be able to conduct research appropriately including review of the literature, methodology, results, and interpretation.

A. First Means of Assessment/Criteria for Success: In Meth 442, our required capstone class, all graduating seniors (70 out of 70) submitted a 10-15 page mini-study that contains a literature review, a methods section, results, and interpretation and discussion. Because this assignment counts for 20% of the final course grade, students take the project seriously. Teams of two faculty members evaluated each mini-study using a locally designed rubric. Using the rubric, each component of the above outcome - literature, methodology, results, and interpretation - was rated on a five-point scale ranging from (1) Beginning, (2) Developing, (3) Meets Minimum Expectations, (4) Exceeds Expectations, and (5) Superior. Each of these scale points was "anchored" by behavioral descriptors to assist the raters in determining a score. In addition, raters went through an hour-long training session to improve inter-rater reliability. Past studies using adjacent-rater reliability revealed 85% agreement across all elements, demonstrating a relatively high degree of consistency between raters. As the criteria for success, the department would like students to average 3.5 or higher across all elements and for each element, a score which translates into somewhat exceeding our minimum expectations.

*Results*

Average Faculty Ratings of Students' Mini-Studies

	Overall	Literature Review	Methodology	Results	Discussion
Seniors 2006 (n = 70)	<b>3.8</b>	<b>3.6</b>	<b>4.0</b>	<b>3.9</b>	<b>3.8</b>
Seniors 2005 (n = 64)	3.6	3.1	3.9	4.0	3.8
Seniors 2004 (n = 73)	3.6	3.2	3.8	3.8	3.7

B. Second Means of Assessment/Criteria for Success: Number of students whose work is accepted at a peer-reviewed regional or national conference or journal. Students are encouraged to submit and, if accepted, present or publish their research. Almost all of these venues require research in the lit review/methodology/results/interpretation format. Students self-report on the Graduate Exit survey if they had any research work accepted, and the number of presentations they had given. In the past several years, participation has held steady at about 50%. Given that this rate is higher than the national average for similar programs, we are satisfied with a rate of 50% as our criteria for success.

## Results

Participation of Students in Scholarly Activities

	Percentage of students
Seniors 2006 (n = 70)	58%
Seniors 2005 (n = 64)	62%
Seniors 2004 (n = 73)	52%

### Interpretation and Use of Results

On all accounts, students met or surpassed our criteria for success. Students' mini-studies were rated above 3.5 in every area and participation in scholarly work was again higher than 50%. One area that the department is particularly proud of is the 3.6 average rating for the literature review section. Over the last few years, this area has been a relative weakness – students consistently being rated below 3.5. In response, the department handed out examples of good student literature reviews and a sheet of “pitfalls to avoid” to students before they prepared their mini-studies. We believe these interventions have helped students write better literature reviews. We are satisfied with the results and will stay the course for now.

**III. Intended Learning Outcome 3:** Seniors will be able to evaluate social science articles, specifically identifying areas of strengths and weaknesses in each of these major sections: the lit review, methodology, results, and discussion.

A. First Means of Assessment/Criteria for Success: Two or more faculty members used a rubric to evaluate each seniors' critique of an article; each student selects his or her own article. The assignment was submitted by all graduating seniors (n = 70) as part of their capstone class, Meth 442. Because this assignment is 20% of their grade, students put forth good effort. The rubric used to assess the critique was originally created by the University of Oklahoma's Measurement department. It was selected as an evaluation instrument because (1) it covers all of the topics stated in this learning outcome, (2) created by group of experts, and (3) provides behavioral anchors that aid with rating. Each item was rated on a four-point scale, from poor to excellent. Considering the rubric, the department decided that an average score of “3” or “good” across all traits and for each trait would indicate success.

#### *Results:*

Average Faculty Ratings of Students' Critiques

	Overall	Literature Review	Methodology	Results	Discussion
Seniors 2006 (n = 70)	<b>2.9</b>	<b>3.0</b>	<b>2.6</b>	<b>3.1</b>	<b>3.0</b>
Seniors 2005 (n = 64)	3.0	2.9	2.6	3.3	3.0

B. Second Means of Assessment/Criteria for Success: On the exit interview (details already given), all seniors (n = 70) indicated how the department can make them more competent in evaluating the methodology sections of articles. In the past, the faculty's review of students' critiques has revealed relative weakness in evaluating the methodology section. This question is meant to probe students about this problem in the hopes that their suggestions may inform future curriculum planning. Criteria for success are not applicable to this assessment.

*Results:*

As is often the case with open-ended questions, answers varied considerably. Nevertheless, two main themes emerged from students' comments, which I will summarize here.

1. That section is the most complicated.
2. Many aspects of the Methods sections of the articles were more advanced than anything students had read in class.

Use and Interpretation of Results:

The results of this assessment are mixed. The overall average rating of 2.9 for the critiques is slightly below our criterion for success of "3" or "good." The culprit for this sub-three rating is the relatively low score of 2.6 on the methodology section. All other areas met our expectations. This weakness in methodology is worrisome because this area is emphasized heavily in the major.

Faculty considered these results in a department meeting and decided that perhaps the problem was in the assessment. Because students were free to choose which article they picked, many students selected articles with methodologies far more advanced than what they've encountered in coursework. Consequently, they did not critique that section of the article well and were rated lower. From an assessment perspective, the faculty member's rating of students' critiques was not only influenced by students' proficiency but also by the difficulty of the article. In other words, letting students pick their own articles introduced a major confound to the assessment. To correct this flaw, faculty decided to select one article that all students will critique thus eliminating the confound. We hope that this change will make our assessment of the critiques more meaningful and useful in the future.