Selecting which outcome and approach

In fall of 2009, the English Department faculty developed an assessment plan for student writing in Composition II classes. A task force established a benchmark for outcomes by examining a sample of Composition I papers and then developed an evaluation rubric to be used for Composition II papers. The task force then collected sample Composition II papers in spring 2010 that were scored during that summer. Weaknesses observed in student writing samples led to a redesign of both courses, and alternative approaches to developmental versions of Composition I for less prepared students.

The faculty concluded that the need for continued assessment was so great that Core Advisory Council (CAC) began a writing assessment project in 2012 to apply the department-developed rubric to a wider sample of student writing to assess overall UIW student learning in this core outcome. The latest outcome statement from the Undergraduate Bulletin reads:

> …[students will be able] to write and speak clearly and persuasively and to convey meaning effectively in non-verbal contexts.

The use of rubrics to score student writing had proven very effective in the 2009 study as a means to evaluate student writing in four broad areas: Purpose, Support, Organization, Style/Form. However, the overall approach for this core outcome assessment needed to measure student writing in courses other than English, and the department faculty recognized the need to adjust the rubric to make it more instructor-friendly for faculty in all disciplines. Rubric evaluation blocks, for instance, had guidance and vocabulary appropriate for English-trained faculty, while those instructions might not be as clear, for example, to science faculty.

As a result, the writing assessment project began with a workshop (January 2013) for faculty from multiple disciplines, to introduce them to the rubric, to train participants to score papers consistently, and to prepare faculty to use the rubric in their own sections during the semester. The workshop was organized to help assure inter-rater reliability and to orient the graders to the rich research already accomplished by the English department [see Minutes – CAC – 12-17-2012 for workshop planning notes].

The rubric we started with

In the English department's 2009-2010 writing assessments, they began with the AACU VALUE Rubric for Written Communication, and modified it slightly for department use. The UIW rubric underwent additional revision through that study, and was clarified for use by non-English departments for use in this 2012-2013 core assessment project. The final rubric and its companion explanatory narrative are separate attachments; the rubric still has the same basic structure as in the 2009 study, with four evaluation areas of Purpose, Support, Organization, and Style/Form.

Writing assessment workshop and instructions to participating faculty

In Fall 2012, invitations went out to all full-time faculty, and many part-time faculty teaching core courses, to a workshop on how to use rubrics to assess student writing. The core assessment team made a strong effort to recruit faculty from the adult and online programs to be sure the assessment would allow comparisons of students in courses with multiple modes of delivery [see the attached Invite to Participate – to EAP faculty].

Invitations pointed out that the study would measure the general writing competencies of our current undergraduate students, irrespective of their chosen academic major or the primary modality of course instruction. Participation in the study required each instructor to choose one writing assignment already planned for the term. Preferably, this would be a 3-5 page final draft of an assignment that students would work on mostly outside of the classroom. However, lengthier pieces were not immediately excluded as viable candidates.

In their classroom implementation, instructors were coached to explain to their students the general purpose of the pilot study and then share the attached rubric with students before they were asked to complete the assignment. Using the four sets of criteria found on the rubric, participating instructors would evaluate each essay submitted by students enrolled in his/her sections.

The evaluations and samples, including a description of the assignment itself, were then be submitted to the core assessment team at the end of the term for further processing and analysis. In particular, faculty were asked to submit at least three scored papers, with low, mid-range, and high scores; in many cases, faculty provided more samples than that, and sometimes the entire section of scored papers. The names of students and other unique identifiers such as identification numbers were redacted.

Courses and assignments collected

A total of 36 main campus samples and 89 Extended Academic Programs (EAP, adult and online division) samples were scored. Main campus samples came from sections of World Literature, World History II, US History II, Marketing, Literary Theory, History of SE Asia, Texas History, and History of Colonial America. EAP sections included: Marketing, Management Theory, US History I & II, Social Studies Instruction, Composition I & II, International Entrepreneurship, and Introduction to Philosophy. These sections were all taught by faculty who had attended the rubric workshop, or had one-on-one preparation sessions with members of the core assessment team (augmented by an instructional video on use of the rubric http://screencast.com/t/5GXvJX89NB9).

Instructors graded their samples first, using the rubric, then turned in the scored papers and rubrics to the core assessment team.

Data analysis and findings

Members of the core assessment team, David Stein (assessment director for EAP) and Karin Simelaro (adult program faculty member), analyzed the aggregated rubric scores from all the participating faculty members.

Inter-rater reliability. Their analysis began with spot-checking of the scored samples that participating faculty submitted. David and Karin rescored 25 essays sampled from all the participating sections, and found that inter-rater reliability was very strong, with 88% pairs of scores identical, and all the remaining 12% were no different than one level.

Validity. We relied on the many years of VALUE rubric assessment – the basis for our slightly modified rubric – for the validity of the instrument to genuinely measure student learning.

Statistics.

Faculty were invited to participate from a wide variety of courses to allow comparisons of student writing skills. The core assessment team received more scored samples from EAP sections than from main campus sections; this strengthened the statistical significance of the findings, but did not weaken the analysis.

Statistics from EAP sections, for instance, showed significantly better writing for upper-division students compared to lower-division sections, in all four evaluation areas. Somewhat surprisingly, adult learners (ADCaP) scored measurably better than online students.

		Purpose		**Support**		**Organization**		**Style/Form**	
Lower Level	n=39	3.41	***	3.23	**	3.33	*	2.85	**
Upper Level	n=50	4.02	***	3.70	**	3.72	*	3.32	**
ADCaP	n=49	4.08	***	3.71	**	3.71	*	3.27	
Online	n=40	3.35	***	3.23	**	3.35	*	2.93	
* Statistically significant at alpha of 10%									
** Statistically significant at alpha of 5%									
*** Statistically significant at alpha of 1%									

Similar trends were observed for main campus sections where upper-division students showed stronger skills in all four areas, compared to lower-division students. The statistical significance was not as strong, because the sample sizes were relatively small.

		Purpose		**Support**		**Organization**		**Style/Form**	
Lower Level	n=22	3.45	*	3.44		3.43		3.10	*
Upper Level	n=14	4.21	*	3.70		3.95		3.94	*

Comparing EAP and main campus students in identical courses, we found that student writing skills were virtually indistinguishable. In the overall data in the tables above, the shaded blocks show stronger student writing scores in upper-division main campus students, compared to EAP results. However, these comparisons might not be statistically significant due to the smaller sample size for main campus sections.

Final rubric and recommendations

Approximately 200 student papers were scored using the common rubric. The scoring statistics allowed CAC to draw interesting conclusions from the data results:

- Upper-division writing scored higher than lower-division for both main campus and students in Extended Academic Programs (EAP; including ADCaP and UIW Online)
- The Style/Form category was the weakest, and Purpose the highest, for both the main campus and EAP.
- In EAP, online writing was weaker than in ADCaP.

Faculty generally thought the rubric was adaptable for use in any discipline.

These results were very timely, as all these five years of direct assessment of student writing feed directly as baseline data and priorities for UIW's 2015 QEP to improve student writing.