

Easton Paper

Professor Celia Easton, one of the designers of INTD 105, presented this paper on its creation and early assessment at a March 2001 SUNY writing conference.

To go back to the main page, click [here](#).

Celia A. Easton
Department of English
State University of New York College at Geneseo
Geneseo, NY 14454

(paper presented at SUNY Conference on Writing, March 10, 2001, Alfred State College)

Assessment of Geneseo's Writing Seminar

Abstract

This paper describes Geneseo's new Writing Seminar and its assessment. I begin with the planning process for assessment for SUNY Geneseo's Writing Seminar. I describe two aspects of the assessment implementation: a survey of students' self-evaluation in terms of the course's learning outcomes, and a scoring of a sampling of anonymous student papers according to a writing rubric. I draw some preliminary conclusions about the Critical Writing and Reading Core after one semester of teaching and assessment and make recommendations for future assessment and faculty development.

Background to the Writing Seminar

Students entering SUNY Geneseo in the Fall 2000 semester were the first class required to satisfy the new Critical Writing and Reading Core requirement by taking INTD 105 Writing Seminar: (subtitle). This course and Core area were created after a period of General Education curriculum review that produced these results: faculty wanted to change the current Critical Reasoning Core requirement, which could be satisfied with either mathematical or written analysis, but did not assure that all students would experience both skills; and students wanted a course that focused more particularly on writing skills. The Writing Seminar is taught by faculty across the curriculum, including the departments of Art History, Biology, Chemistry, Education, English, Geography, History, Music, Philosophy, Political Science, Psychology, School of Business, and Sociology. Enrollment was originally set at 25 students per section, but widespread interest by faculty willing to teach the course has enabled the College to reduce class size to 22. Students must elect the course during the fall or spring semester of their freshman year. Approximately fifty sections are offered each year.

The College chose to have the course taught by faculty across the curriculum to emphasize to students that critical reading skills are important for all academic areas. Faculty offer students sections of the Writing Seminar focused on discipline-specific topics, such as "Evolution," "*The Scarlet Letter*," "The Philosophy of Love," "Twentieth-Century Composers," and "Tocqueville's *Democracy in America*." Each section is directed to require students to explore significant texts slowly and analytically. The Writing Seminar is not intended to be a course that teaches a topic broadly; rather, students are expected to study a few works in depth.

Sections of the writing seminar share one common requirement: students must write six assignments for a total of approximately twenty pages during the course of the semester. There is no mandate about the nature or length of individual assignments. The six assignments may include revision of papers, if revision is taught as a distinct writing task. We have made room for diverse approaches in regard to the actual teaching of writing. Many faculty make their students' papers central texts in their courses; others devote class time to talking about the writing process and particular features of writing essays. But no faculty member is technically required to do more than conduct a discussion course on a particular topic and assign six papers to the students.

The Core Committee received some criticism from experienced teachers of writing for allowing so much instructor autonomy. By leaving the pedagogical task open, the Core Committee hoped to encourage widespread faculty participation in the teaching of INTD 105. More than half of the faculty who taught the course during the Fall 2000 semester had never taught a writing-focused course before. At a post-semester review lunch, several noted that when they taught the course again, they planned to restructure the time they devoted specifically to writing instruction. Undoubtedly, some will not arrive at this insight on their own, and some students will study in sections where writing instruction is not stressed. The Core Committee's position has been, however, that students learn to write by writing, and they will nevertheless be well served by the course. One tool for writing shared across all sections of INTD 105 is Geneseo's Online Writing Guide (<http://writingguide.geneseo.edu>). While reminding students that all writing is conventional, The Guide also provides uniform criteria for good writing expected by professors at Geneseo.

The College provided a week-long pedagogy workshop for INTD 105 instructors at the end of the Spring 2000 semester. The Provost encouraged attendance by giving workshop participants \$500 stipends if they took the workshop and taught a section of the course. Various workshop leaders, also paid by the Provost, gave presentations on responding to students' writing, teaching revision, teaching critical reading, conducting peer editing sessions, and addressing issues of diversity in the Writing Seminar.

We have now begun the second semester offering of the new writing seminar. Because of the quantity of faculty interest in offering sections of the course, we have been able to reduce class size to 22 students, guaranteeing each freshman at least one course in which they can develop a significant relationship with a faculty member and all other class members. During the January gathering of Writing Seminar professors, many who offered Fall 2000 sections cited these relationships as among the most rewarding aspects of the course for faculty and students. This is a reminder that many course "outcomes" that contribute to students' total educational experiences are never articulated in formal assessment procedures. They are, however, often more meaningful to our students than the academic skills we seek to measure in assessment.

The Critical Writing and Reading Core at SUNY Geneseo is unique because its plan for assessment was created at the time the Core was instituted. Motivation for assessment was the best kind. This is a new course and new core area, so faculty were willing to ask, "What can we do to make it successful?" While assessment plans are generally formed in response to mandates of local administration, SUNY Trustees, or state legislators, the assessment plan for the INTD 105 Writing Seminar was also motivated by genuine interest in improving student learning at the College.

Assessment Planning

The Core Committee, which I chaired during the 1999-2000 academic year, created the assessment plan. Having attended a number of assessment workshops, I created a short guide to assessment for the committee, which admittedly contained a number of straw men. [appendix 1] We were not going to administer standardized tests, and we needed significantly more information than students' opinions about their success as writers. We lacked the financial resources to make anthropological site visits and qualitative reports. Our focus was on **program** assessment, not the assessment of individual students or instructors.

The first task was to formulate intended learning outcomes for the course. Given the autonomy the course allows individual sections, the outcomes needed to be basic enough to fit each student's experience. We wrote three outcomes, which are reproduced on each section's syllabus. Two address writing; one addresses reading. [assessment plan, appendix 2]

Implementing Assessment: survey on learning outcomes

To accumulate data on reading and writing skills, we scored a selection of papers from students enrolled in the course according to a writing rubric. We supplemented that data, however, with a survey of students' opinions about how well the learning outcomes were met by the course. [appendix 3] Five hundred fifty-eight students responded to the survey. On the whole, students scored the course positively, with over 60% indicating that they strongly agreed that the course improved their writing skills; 57% claimed their reading skills improved significantly because of the course.

Such surveys provide limited assessment data, but they do help to promote good will about the course. Many students have a pre-determined resentment toward general education requirements. Students will attend classes and complete course work more consistently, however, if they believe that the course has a genuine academic value. Thus far, student satisfaction with the course seems to match faculty enthusiasm.

Comments written on the student assessment surveys do raise a few areas of concern. Although all faculty believe that they are teaching reading skills by modeling close reading and analysis, we frequently fail to explain what reading "critically and carefully" means. Thus, many students think of "reading skills" as comprehension, vocabulary building, and speed—the kinds of skills emphasized by primary and middle school teachers. Many college students are concerned by their inability to read quickly, even when faculty try to emphasize quality of analysis over quantity of pages.

When we ask students to think about "improving" as writers, we need to remember that they focus more on the grades they receive than the comments written on papers as marks of their skill levels. Thus many students reflected on their surveys that they believed their writing skills improved, but they had received the same grade—a B—throughout the term. Finally, we need to remind ourselves that words we think of as pedagogic jargon have different meanings for our students. While most writing teachers agree that all kinds of writing assignments have "arguments," some students retain a negative connotation of the word from high school teachers who tell them not to write their "opinions." Thus one student wrote in response to the second learning outcome ("ability to write sustained, coherent arguments"), *"My argument was less meaningful, allowing me to concentrate on the actual writing."*

Implementing Assessment: scoring essays with a rubric

The larger part of the assessment process for the Writing Seminar involves scoring a selection of papers from each section of the course according to a writing rubric. [appendix 4] The rubric, created by the Core Committee with input from several members of the English Department, emphasizes thinking skills in the categories of "argument," "logic," "evidence," and "directness." One category addresses reading; assessment volunteers found that category most difficult to judge from the essays they read. The final column on the rubric addresses mechanical and grammatical errors. Isolating predominantly surface errors to one sixth of the rubric was an attempt to remind professors and students that writing involved much more than spelling, punctuation, and proper documentation style.

In each section of the Fall 2000 Writing Seminar, students turned in extra copies of their fifth or sixth class assignment. A secretary in the English Department made sure that papers were anonymous. Since we did not have the resources to read all of the 550 papers we collected, I turned to two statisticians to determine what a legitimate sampling would be. A professor in the School of Business came up with a number that was close to 70%; the Sociology Department's statistics expert suggested a number closer to 40%. I sent their emails to the Associate Provost who oversees assessment in the College, who also happens to be a Sociologist. Since two assessment volunteers would read each essay, I doubted that we would be able to score such a large number. He told me that we could not expect statistical accuracy and advised—I hope not arbitrarily—that we read fifty essays.

The 2000-2001 Chair of the Critical Writing and Reading Core Committee, Professor Rachel Hall, randomly selected two essays from each section of the course to be read by teams of faculty members. Each team read approximately ten essays; two readers scored each essay. A secretary in the English Department entered the one hundred scores into an Excel document, identifying the papers only by number. [appendix 5] We do learn something about students' writing from these scores. We also learn something about the assessment process, and perhaps about ourselves as readers of student writing.

The rubric attempts to distinguish particular information about students' writing, without making judgments about how much each category counts in a paper's total success. Visually, however, six equally sized columns imply equivalent importance. Even though this was not a grading session, some assessment volunteers found it difficult to ignore the rubric's graphical suggestion that each column represented one sixth of a grade.

Many of the assessment volunteers noted after the process that they should have spent more time as a group discussing what the rubric meant and how categories should be judged. It is easy to count misspellings and dangling participles. Scoring "directness" in writing, however, is far more subjective. We agreed to use the rubric through another round of scoring at the end of the Spring 2001 semester before considering how we might wish to refine it as an assessment tool.

Before this round of assessments, I received some comments from colleagues who described students who wrote good papers but were sloppy spellers. That is precisely the reason for separating out "grammatical clarity" from argument, logic, etc. "Grammatical clarity" covers a large number of surface errors a writer might make, from misspellings to fragment sentences. The expectation of competence is fairly strict: on a scale of 1 to 5, a student must have no more than four errors in sentence structure, spelling, word choice, punctuation, or documentation to receive a score of "3." When the reading team's scores were averaged, approximately half the student essays fell below a "3" in the grammatical clarity column.

One surprising result, however, was that in comparison with the other data for the rubric columns, the "grammar and mechanics" column is not extraordinarily low. Student averages for "grammar" were 2.67. They had the same averages for "argument" and their averages for "logic" were even lower: 2.54.

The correlation between argument and logic and students' spelling, punctuation, and grammatical flaws suggests several possibilities. Some writers who do not pay attention to surface errors also neglect steps in their arguments, transitions between paragraphs, and sufficient evidence for their claims because they are hurried writers whose written output gets ahead of their thinking. It is also possible, however, that graders get so annoyed by the surface errors that they score students' performance in constructing arguments more harshly when their reading is interrupted by inattention to spelling, punctuation, and grammatical convention. For fifty papers we had 100 grade reports. Forty-six grade reports scored students as less than competent (1 or 2) in grammar and mechanics. Of those 46, thirteen were scored competent (3) for their arguments, and four were given scores higher than competency.

On the whole, students who performed grammatical and mechanical skills well—above a 3—did well in other areas of writing performance. Of the 100 grade reports, twenty-three scores of 4 or 5 were given for grammatical and mechanical skills. Of those 23 students, six had scores lower than 3 (competency) in one or more other columns.

How accurate are the data gathered by scoring papers according to a writing rubric? In most cases, pairs of assessment volunteers scored papers within a point of each other. While this may not seem to be a significant difference, in some cases one point distinguishes a "competent" from a "less than competent" writer. I have already noted the volunteers' desires for additional discussion about the rubric categories before the scoring sessions. But we must also bear in mind that however much numerical scoring produces data that appear to be objective, the process is very subjective. One example of this is the scoring for the category "evidence." Student paper #15 received scores of 1 and 3.5 from one pair of assessment volunteers. The paper addresses the growth of a character in *War and Peace*. Throughout the paper, the student makes claims about the character, then draws on examples from the novel to back up the claims, as in these two sentences from paragraph two: "When we first meet Andrew we find first and foremost his distaste for aristocratic society. Tol'stoi uses adjectives like 'weary,' 'bored,' and 'tiresome' to describe Andrew's disposition at Anna Pavlovna Scherer's soiree." Although the writer fails to develop the analysis of these examples, he or she follows this pattern throughout the essay, matching assertions with textual evidence. The student even draws attention to evidence on the third page of the essay by saying, "Perhaps the strongest evidence [for Tol'stoi's interest in this character's development] comes from Tol'stoi himself in a letter to L. I. Volkonkaya on May 3, 1865...." Why, then, does one reader conclude that this writer uses evidence incompetently? Faculty readers do not, in fact, agree about the meaning of the terms of the rubric. Some faculty members penalize students who gather evidence but fail to draw significant conclusions from it. Other faculty associate "evidence" with secondary research, while still others tell their students that "evidence" means supporting their assertions with concrete examples, no matter the source.

Preliminary Conclusions

Rather than enumerate each category and its potential for discrepancies, I want to address a broader assessment question: What did we learn from this process that could be of use to improve student learning? I have been asked by several faculty members why we are not assessing writing competency by reading papers written at the beginning of the term and comparing them with papers written at the end of the term. My answer is that we are not assessing individual students' development or writing competency. Rather, we are looking at student writing from the end of the term to determine which areas we need to strengthen in the course. Because we have gathered data from only one semester, these conclusions are preliminary, but they provide us with direction for ongoing faculty development. At this time I want to focus on two areas:

1. **Logic and Coherence.** 50% of the readers' reports scored students' essays as less than competent (1 or 2) in "logic and coherence." This means that essays were judged faulty in organization, in ability to draw inferences, and in focusing on the topic. We will address this explicitly at the next faculty pedagogy workshop, and develop written materials for those who cannot attend to encourage instructors of the Writing Seminar to address this area with particular care in their courses.
2. **Critical Reading.** The Writing Seminar is a text-focused course, although the main course text might not be a written text. While most students received high marks in Critical Reading during this round of assessment scoring, many assessment volunteers questioned their own ability to judge students' insights about texts the volunteers had not read. The current assessment vehicle is indeed inadequate for studying students' engagement with the texts in their courses. The student opinion surveys on intended learning outcomes, however, suggest that students need more information about what it means to read critically in college. Although we did address the teaching of critical reading in our faculty development workshop last spring, I believe that we must emphasize to instructors the need to go beyond modeling the reading process. While this course was not developed as a "study skills" course, students would be well served by classroom comparisons of reading for different purposes and discussions about developing insightful, informed observations about texts.

Recommendations for the Future

How should we modify the assessment process? The scoring of the sampling of essays proved to be a fairly simple task that took approximately three hours of ten volunteers' time. The data were compiled by one secretary and one work study student in the English Department. We could improve the accuracy of the data by providing more training time for assessment volunteers and increasing the sample size. Doing so, however, increases faculty members' time commitments and risks the good will of the current pool of volunteers.

The scoring rubric has achieved some popularity with faculty not involved in the Writing Seminar as a resource for articulating to students which areas of their writing need improvement. To function better as an assessment tool, however, we need to have greater ownership of the rubric's concepts about writing competency. This might mean changing the rubric's categories and criteria to reflect expectations about writing held by faculty across the College. I anticipate that we will hold a workshop on revising the writing rubric next year.

With a refined assessment rubric, however, we will be able to move the assessment process away from a fifty-paper sampling and into the hands of individual instructors. If half the faculty teaching the Writing Seminar each semester scored two papers per term according to the grading rubric and submitted the data collected to the Core Committee, we would have a large and more revealing data pool. The accuracy of the data would be particularly improved in the category of Critical Reading, since instructors would determine their own students' engagement with the course texts. Instructors would also have better insight into the expectations of a given assignment, and whether or not their students were, for example, using evidence as they had been instructed.

Assessment plans that use instructor scoring for data collection eliminate the redundancy of re-reading papers that have already been evaluated by the course's professor. The drawbacks of instructor scoring are significant, however. If instructors fear that their students' performance will be considered a reflection of their own teaching performance, they may distort the assessment data. We need to reassure faculty that the purpose of the assessment is to evaluate the program, not instructors' or particular students' competencies. A great deal more training will be needed to get all instructors comfortable grading with a rubric. We need to pay careful attention to the feedback from the first group of assessment volunteers who wanted a greater sense of consensus on the terms of the rubric. The good news, however, is that instructors who already incorporate a rubric into their grading process find that it actually saves them time because they do not have to write out comments already addressed by the rubric categories when they return papers to their students. The categories also remind students that an assignment's grade is determined by a large number of competencies.

Instructor-based data collection cannot be implemented without considerable faculty training. For the Spring 2001 semester, we will repeat the sampling process we followed for last fall's sections and take another survey of students' attitudes about the learning outcomes. The process has made us all keenly aware of how new this Core area is, and we are eager to use the flexibility of that newness to make improvements in the course that will benefit all our students.

To go back to the main page, click [here](#).