
Writing Assessment at Plymouth State College

by Russell Lord

The Nature and Purpose of Assessment

Traditionally we associate the term assessment with appraisal of real estate or other property. An assessor attempts to set a market value upon a particular object by comparison with similar property. It is thus an attempt to evaluate something for market purposes. When we attempt to apply the term to academic pursuits, we tend to transfer these connotations to the academic community, where there is a developing, and some believe destructive, trend toward applying the idols of the marketplace to the free pursuit of knowledge, where they do not belong. And yet a college, especially a state college, must set a nice balance between its accountability to the public, which gives it support, and those professors who remain committed to the ideals of the most effective teaching as they envisage it. Measuring the results of such teaching is as complex as the teaching process itself; and, as our experiment proved, its results are by no means so dramatic as the marketplace tends to demand.

Background

In the spring of 1989 the Writing Across the Curriculum program at Plymouth State College had been established and functioning for approximately four years. The Dean of the College (Theo Kalikow) and a group of faculty chaired by Sally Boland determined to judge whether the college curriculum improved student writing during students' four-year college experience. Pressures from the legislature in New Hampshire and, indeed, throughout the nation demanded some type of evaluation procedure, and this committee

bravely set out to avoid the pitfalls of too mechanical an assessment, dependent on an objective test or two which could give neat statistical results but which would fail to address the complexity of the writing process itself. The committee devised a combination of an essay test graded outside the college and portfolios evaluated by our own faculty. The essay tests would be given at the beginning and end of the first-year Composition course, and then during the spring semester of the fourth year. The portfolios would contain written material primarily from general education classes gathered over four years of the students' college experience.

The Instrument

For the outside-graded portion of the assessment, the committee ultimately chose the essay test provided by the American College Testing Bureau (ACT) in Kansas City, Missouri, a segment of the Collegiate Assessment of Academic Proficiency (CAAP) testing program. The test consists of two 20-minute writing samples requiring students to support a position on a clearly defined issue. In the Fall of 1989 the test was graded on a four-point scale. Currently ACT uses a six-point grading system, but the company agreed to return to the four-point system for our Spring 1993 tests so that our results might be consistent. The criteria for the four-point scale follow:

4--Substantially developed appropriate argument. These papers take a position on the issue defined in the prompt and support that position with an argument of one or more appropriate reasons. The argument's main ideas are logically connected and substantially developed.

3--Moderately developed appropriate argument. These papers take a position on the issue defined in the prompt and support that position with an argument of one or more appro-

priate reasons. The argument's main ideas are logically connected and one or two may be moderately developed, but the argument as a whole does not constitute an elaborated argument.

2--Minimally developed appropriate argument. These papers take a position on the issue defined in the prompt and support that position with a brief argument of either two or three appropriate but undeveloped reasons, or one appropriate reason only minimally developed. These papers recognize the grounds upon which the issue will be resolved, but the argument does not focus on those grounds.

1--Insufficient or inappropriate argument. These papers take a position on the issue defined in the prompt but offer only one undeveloped appropriate reason in support of that position. Or these papers take a position but do not support that position with any appropriate reasons.

The portfolios in that segment of the study graded by our own faculty consisted of materials ranging from term papers or other student essays, to lab reports and essays on examination questions. The courses involved were primarily drawn from those in the General Education program at Plymouth State College, although other courses with adequate written responses were also included over seven semesters of the students' college experience. Grading by our faculty used a holistic scale with criteria that we decided should include the following elements:

- Quality of thought. This included a student's depth of understanding of the problem involved and ability to convey that understanding to a reader.
- Quality of expression. This involved the organization of the materials in a coherent mode that showed an understanding of rhetorical principles fitted to the nature of the

essay involved.

- **Mechanics.** These involved a mastery of sentence structure, spelling, punctuation, and diction appropriate to the writing situation.

We used a six-point scale to evaluate these elements, placing more weight on the first two criteria than on the third.

Development and Characteristics of the Instrument

Members of the English and Education Departments, with cooperation of the Dean of the College, devised the instrument. They decided to draw a random sampling of six first-year Composition classes to be divided into three groups: two classes to serve as a control group to be assessed only by the outside graders; two classes for whom portfolios of significant writing would be maintained for four years; and two classes for whom portfolios would be maintained but who in addition would be coached in their writing for the four years. Since each class had approximately 25 members, our study would cover initially about 150 students. Each student signed a form, indicating willingness to take part in the study, with the opportunity to withdraw at any time from it. Five Composition teachers chose to be involved (Sally Boland, Arthur Fried, Mary Lou Hinman, Walter Tatara, and Gerald Zinfon) and Russell Lord became Director of the project.

Our first need was to devise a way to gather and house the portfolios. Peter Hart from the Computer Center, in cooperation with Bill Clark of the Registrar's Office, devised a computer program to store the names of designated students and requested their instructors to send copies of papers for the portfolio. The papers were kept in manilla folders maintained in a file cabinet in my office.

During the second year of the study, we held a workshop

directed by Sharyn Lowenstein from the writing center at UNH Manchester. We discovered that a holistic approach with loosely defined elements yielded a much greater degree of consensus than a method using a rigorous series of weighted criteria, and this experience formed the model for our later grading of the portfolios themselves.

As is almost inevitable in such a study, a serious problem evolved during this same year. The students who were to be coached during their college experience did not desire that benefit. We were therefore forced, for this pilot project, to restrict the portfolio study to the effect of the normal college program itself on writing progress over four years.

Because of attrition we ended the study with 43 of the original 85 usable portfolios; 19 of these students took the CAAP test in their Senior year. In addition, seven from approximately 27 students remaining in the control group took this CAAP test.

Five faculty members did the Portfolio grading: three English Professors (Mary Lou Hinman, Arthur Fried, and Russell Lord), one Psychology Professor (Robert Miller), and one Chemistry Professor (Wavell Fogelman). Two readers examined each paper, whenever possible representing two disciplines, and through discussion arrived at consensus. Student names, dates, and grades were removed from each portfolio essay to be evaluated.

Statistical Results of the Study

Robert Hayden of the Mathematics Department, in a statistical analysis of the results, made essentially the following observations:

I tried many multiple regression models to see how portfolio grades were affected by the other variables. Only two variables were consistently important: although Portfolio grades generally tended to rise by about 0.1 point over the

period, a student who received above C on the English Composition grade generally had a portfolio grade about 0.7 points higher at the end of the 7th Semester than at the beginning of the college experience. There was thus some evidence that students with A's and B's in English Composition showed more improvement over time than those with C's or below.

I also found CAAP scores and the various GPA measures to be interrelated, but this family of measurements was not particularly related to portfolio grades or English Composition grades. Also, the grade received in the course in which a portfolio sample was evaluated was not related to the grade the relevant portfolio essay received.

Conclusions

1. For a project of this type, in order to negate the influence of attrition, a larger initial sampling would be desirable. From our experience we might predict approximately half of the initial sample would remain over four years.

2. A deeper commitment of students to the program needs to be carried through the four years. For that purpose the college needs to offer students greater incentives. We gave refreshments to induce them to attend the CAAP test in their senior year; yet the offer failed to draw many of them. Almost none showed interest in the proffered coaching during their four years. Academic recognition of some type seems necessary if we expect more deeply engaged students. And although the Dean and I wrote several letters to the students during the course of the study, publicity directed toward the students needs to be created.

3. Although statistical results were generally not impressive, comparison of the portfolio essays with Composition grades seems to show that success during the first-year English Composition course influences writing progress over the four years.

4. Statistics support a correlation between the CAAP essay test and the GPA, but not between CAAP and the portfolio or composition grade. This result may indicate that the writing samples are better related to issues of critical thinking than to rhetorical techniques, and are thus more an indication of overall academic progress than of writing per se.

5. Although it was not conceived as a goal of the study, one of the most productive results was its effect on the faculty. It created an interest in student writing as a means of effective expression and analysis of course content, rather than as an almost irrelevant ancillary to the content itself. And to those actively involved in the Assessment process itself, it provided an opportunity to think more effectively about just how to evaluate student papers.

Assessment Steering Committee Conclusions

At its final meeting the Steering Committee concluded from the study the following points:

1. If we were to repeat such a study, we would have to have a much more committed group of student volunteers, with powerful incentives.

2. Portfolios would need tighter control of material, so that uniform contents would yield more measurably consistent results.

3. Providing the assignments for each portfolio essay would be desirable for proper judging of contents.

4. The most telling results might come not from graded essays at all, but from surveys of student attitudes toward writing during their college careers. How the students perceive themselves as writers would form a better indication of the way our process-oriented WAC program is succeeding than the portfolios could possibly reveal.

Was the study worthwhile?

As an indication of the success or failure of the WAC program at Plymouth State College the study has many weaknesses. As a means of creating faculty interest in writing, however, it has had some unexpected success. By gathering and sending materials to be included in portfolios, teachers became actively involved in developing writing awareness. Those involved in grading portfolios gained experience interacting in a process which tends too often to be a private preserve with little oversight. The steering committee gained experience in facing unforeseen problems and in solving them in novel ways. The study was thus valuable in revealing need for much more thorough accounting of student motivations and heeding them.

The most significant statistical result of the study, a greater improvement in writing over four years for those doing well in their First Year Composition course, might provide the impetus for fine tuning that course to yield the incentives for greater student interest in their own writing. Current plans to establish a WAC writing center would then provide a means to carry on that interest through the student's college career.

(Note: Data and statistical workings of this study are available upon request.)