In an attempt to contribute to the development of students' writing skills while helping them master basic psychology, a writing component was designed for a medium sized introductory psychology course. Its purpose was to aid in the mastery of course material and to contribute to the development of students' skill in written expression. In the first of two studies, 84 students wrote weekly essays and were graded in part for their ability to write complete sentences and use reasonable grammar. The second study restricted the types of essay questions to four classic rhetorical forms studied in freshman composition. Eighty-one students were given examples of how to write answers to such questions and were no longer graded explicitly on their complete sentences and reasonable grammar; instead they were given a score based on the organization and flow of their essays. Results demonstrate that students do not necessarily improve in their writing skills merely because they are required to write and are being graded on their efforts. Instead an improvement in writing skills was established only after narrowing the kinds of essays students wrote, providing suggestions about writing such essays, and grading the students on the overall quality of their writing. (Three tables giving examples of essay questions and data are included and seven references are appended.) (MS)
Written Assignments in the Introductory Course

R. J. Madigan, James Brosamer, and Sonya Plourde
University of Alaska, Anchorage

ABSTRACT

This paper describes a writing component for a medium sized introductory course. It was designed both to aid in the mastery of course material and to contribute to the development of students' skill in written expression. The paper reports the results of this program on the students' writing skills. In the first study, students wrote weekly essays and were graded in part for their ability to write complete sentences and use reasonable grammar. Independent raters judged the quality of the essays early and late in the semester. No improvement was detected. The second study restricted the types of essay questions to four classic rhetorical forms studied in freshman composition. Students were given examples of how to write answers to such questions and were no longer graded explicitly on their complete sentences and reasonable grammar; instead they were given a score based on the organization and flow of their essays. Once again independent raters judged the quality of essays written early and late in the semesters. The ratings showed a significant improvement. The study shows that students do not necessarily improve in their writing skills merely because they are required to write and are being graded on their efforts. We were able to show an improvement in writing skills only after narrowing the kinds of essays students wrote, providing suggestions about writing such essays, and grading the students on the overall quality of their writing.

Address correspondence to R. Madigan, Department of Psychology, University of Alaska, Anchorage, Alaska 99508.
It is widely accepted that a college curriculum should address the teaching of writing (White, 1985), and there are good reasons for including in introductory courses a requirement for written work by students. Some theorists, such as Block (1985) and Fulwiler (1982), claim that writing contributes to the development of critical thinking; and in fact there is a high correlation between writing skills and certain measures of higher mental functioning such as problem solving (Benton, Kraft, Glover, & Plake, 1984). Requiring students to write can also directly contribute to the mastery of the course material. Several studies (Richards & Friedman, 1978) show that students behave differently toward course material when they expect essay tests. Note taking is more organized and extensive, and studying behavior is more focused when students prepare to write essay tests rather than multiple choice tests. Emig (1977) argues that writing is not merely a mode of learning but a unique mode.

Despite the general acknowledgment of the value of writing requirements, the practical implications of this commitment are not always clear. White (1985) cautions that "we do not know much about the most effective sequencing of composition instruction, nor do we even know if it is possible to develop such sequencing" (p. 246). Although the current interest in writing across the curriculum gives evidence of a renewed commitment to the importance of teaching writing in all
disciplines and although the advocates of writing across the curriculum have had some success in persuading faculty not to hold their colleagues in English solely responsible for the teaching of writing, there is as yet nothing like a consensus on how to share the responsibility. To compound the problem, large enrollments and ready availability of multiple choice and other quickly scored examination devices have made writing assignments seem increasingly burdensome.

For the instructor of an introductory psychology course, then, the added task of teaching writing will seem formidable. A teacher who wishes to serve the ends of general education within the discipline and also to achieve the benefits derived from having the students write must devise assignments that are at least theoretically tenable for teaching composition but that will not require an unacceptable increase in the amount of time spent in reading and grading examinations. Common sense would also dictate that the assignments should be, if possible, supportive of and supported by the school's freshman writing program. We have tried to respond to these challenges in the context of a medium sized (about 130) introductory psychology course. In this paper we report on the results of our attempt to contribute to the development of students' writing skills while helping them master basic psychology.

A Writing Component for Introductory Psychology

Each week during the semester students wrote two essays of about 100 words each on course-related topics. Ten questions
were distributed each week, and in the next class a quiz was administered that contained two of the questions selected at random from the list. Students were encouraged to practice writing their answers before the quiz and were told that part of their grade (five points out of 30) would be awarded for the quality of their writing. It was suggested that part of their preparation for the test should include such prewriting activities as outlining and organizing the material as well as editing or revising trial answers.

A total of 13 weekly tests were administered, each containing two essay questions and eight multiple choice items. The tests were graded by a graduate student in psychology and returned in the next class period. The final exam for the course consisted of six essay questions selected from the set of 26 questions that had appeared on the previous 13 tests.

We have attempted to determine whether this format produced improvements in the students' writing skills by comparing essays the students wrote early in the semester with those they wrote for the final exam. It was our hope that the later essays would show an improvement in written expression. Two variations of this basic format have been evaluated.

Study 1

In Study 1, students took the 13 quizzes and final exam described above. Each week their quizzes were returned with model answers to one of the quiz questions. The
component for this class consisted of the repeated practice opportunities provided by the quizzes, the encouragement provided by the essay scoring system which gave points for writing quality, and model answers that were distributed for some of the essay questions.

**Method**

A total of 94 students registered for the course, and 84 remained registered at the end of the term. Each student received a Study Guide five days before each quiz that contained ten questions which could be answered by a paragraph length essay. Examples are given in Table 1. The quiz in the next class contained two essay questions selected from the two most recent Study Guides. When quizzes were returned in the next class, each essay was graded for complete sentences and reasonable grammar (5 points) and for content (20 points). Legibility was also graded (5 points) to encourage readable handwriting. The subscore given for complete sentences and reasonable grammar was intended to focus students' attention on the quality of their written expression. Each week sample answers to one of the essay questions were distributed. These answers were those judged to be the best in the class by the teaching assistant. Two model
answers to the same question were presented to show different writing styles and approaches.

The data for this study were essays written for the second of the 13 tests and for the final exam. The same question appeared on both exams: Give an example of the use of the experimental method. Students had not received a sample answer for this question.

Useable data were available for 55 of the students. Some students missed either the first test or the final exam. Students taking the final exam could omit one essay question, and some omitted the one critical for this study. Four students were excused from the final on the basis of their high scores during the semester.

The early and late essays for the 55 students were typed onto a single sheet in random order. Four judges read the pairs of essays and selected one as the better of the two according to the six-point scale below.

1: Top essay significantly better
2: Top essay slightly better
3: Top essay, but essays are very close
4: Bottom essay, but essays are very close
5: Bottom essay slightly better
6: Bottom essay significantly better

The judges were two of this paper's authors (RJM and JJB) and two teaching assistants from the English department who had previously taught freshman composition but were unfamiliar with this study.
Results

Table 2 presents the inter-rater agreement for the four raters. The ratings cluster around 0.60. The analysis of writing improvement over the semester was based on the average of the four ratings given to each writing sample. An average formed in this way would be expected to have a reliability of 0.87 (Winer, 1971).

The ratings were adjusted to correct for the order of presentation of the essays. This adjustment produced a scale from 1 to 6 where 1 corresponded to a judgment that the early essay was significantly better and 6 to a judgment that the essay on the final was significantly better. A score of 3.5 would be at the midpoint of the scale where the two essays would be scored to have equal quality. The mean average rating of the four judges was 3.72 (s = 1.5), which is not significantly different from 3.5 [t(54) = 1.07].

Study 1 did not show an improvement in the students' writing over the semester. The essays written at the beginning of the term were judged to be equal in quality to those students wrote in the last week. While the writing may have produced unmeasured benefits for the students in terms of the way they took notes and studied for tests, it did not result in a measurable improvement in the quality of their written expression.
Study 2

In Study 2 we restricted the kinds of essays students wrote in order to focus their efforts on a small number of question formats. We selected the question formats from the basic rhetorical forms stressed in the freshman composition taught in our institution. All essay questions used in Study 2 required one of the following rhetorical patterns: exemplification, definition, comparison and contrast, or process analysis. Table 3 presents examples of each. We also refined the scoring system to give students more direct feedback on writing quality.

-------------------------
Insert Table 3 about here
-------------------------

Special course materials were distributed that discussed each of the rhetorical forms and gave examples of full-credit answers. The students also received sample answers after each quiz. Study 2 examined whether this more focused approach would produce improvements in writing quality over the semester.

Method

The initial enrollment for this class was 88 students, and 81 received grades at the end of the semester. In the first week of the course, students were given a handout with suggestions about how to answer exemplification and definition questions. For the next five weeks, all Study Guide questions were keyed to
one of these two rhetorical forms. In the sixth week questions requiring comparison and contrast were introduced with handout suggestions for writing good answers to such questions. In the ninth week, process analysis was added, and a separate handout gave suggestions about answering this type of question.

The scoring system was modified by replacing the category "complete sentences and reasonable grammar" with the phrase "development" which was defined as the method of approaching the question and organizing content. A five-point scale was used to score development.

Essays were prepared for rating as in Study 1; those of each student from the second week and from the final exam were typed on a single sheet of paper in random order. All essays responded to the question: Use an example of a behavioral study to show that the quality of the early environment is important for neural development. Three judges rated each essay pair using the scale defined in Study 1. The raters were one of the authors (JJB) and the same two English teaching assistants used previously. A total of 36 essay pairs were available for this class. Subject attrition was due to the factors described for Study 1.

Results

The inter-rater agreement for the three judges is given in Table 2 and appears to be close to the agreement of Study 1. Average ratings were calculated for the three judges and corrected for the random order of the essay pairs. The neutral point on the rating scale remained 3.5. A lower rating indicated
a preference for the early essay, a higher rating for the later one. The mean average rating in Study 2 was 4.2 ($s = 1.3$), which was significantly different from 3.5 \[t(35) = 3.16, p < .01\].

Discussion

We have found that it is more difficult to make a measurable improvement in writing than one might expect. It is not sufficient merely to require writing and to encourage students by grading them on their "sentence structure" and "reasonable grammar." This appears not to have been specific enough to be useful to the students in the first study. Even with feedback on their responses, they did not improve significantly. However, there was measurable improvement in Study 2. The writing component was more sharply focused in that the types of essays were restricted to several classic rhetorical forms students encounter in freshman composition. We believe that we had better results because our instructions to the students were more specific. The students were given examples of how to develop their ideas in the context of specific types of essay questions, and the sample answers we gave them showed good development.

Another way of viewing the results is to see the grading system of Study 1 (complete sentences and reasonable grammar) as an interpretation of writing as product and the grading system of Study 2 (development) as an interpretation of it as process. The latter conception is consistent with "development" as an attempt to discover or create meaning by the act of writing. This is one of the central assumptions of advocates of writing across the curriculum. By contrast, the emphasis on sentence structure, on
"reasonable grammar," and, especially, on legibility tended to define writing as editing rather than as thinking.

If we are right in our understanding of the reasons for the modest success of the second study, a likely next step would be to specify further for the student the concept of well-developed writing and to do so with assignments that utilize the common rhetorical modes most appropriate to the content of beginning psychology. Emphasis on development in the specific context of the paragraph should contribute directly to disciplined and orderly presentation of course material and also to the more comprehensive goals of a liberal education.
References


Table 1

Examples of essay questions used in Study 1

1. Why is it said that the experimental method can demonstrate causality but the correlational method cannot?

2. Describe in general the functioning of the autonomic nervous system. Why is this of interest to psychologists?

3. A child's language skill develops in an orderly way. Describe three important milestones in this process and say roughly the age at which each might be expected.

4. Give an example of the use of biofeedback to teach the control of a function regulated by the autonomic nervous system and point out how this is an example of operant conditioning.

5. Using results from animal studies, discuss the effects of deficient early environments on later development.

6. Discuss the social learning view of child behavior problems by describing how children acquire deviant behavior and the approach that social learning therapists take in helping parents of difficult children.

7. Say why reliability and validity are critically important characteristics of psychological tests.

8. Describe some of the cognitive characteristics of moderately depressed persons and relate these to Beck's cognitive therapy for depression.
Table 2

Inter-rater agreements for Study 1 and Study 2 expressed as Pearson correlation coefficients

Study 1

<table>
<thead>
<tr>
<th></th>
<th>Rater 1</th>
<th>Rater 2</th>
<th>Rater 3</th>
<th>Rater 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rater 1</td>
<td>0.57</td>
<td>0.61</td>
<td>0.65</td>
<td></td>
</tr>
<tr>
<td>Rater 2</td>
<td></td>
<td>0.61</td>
<td>0.64</td>
<td></td>
</tr>
<tr>
<td>Rater 3</td>
<td></td>
<td></td>
<td>0.65</td>
<td></td>
</tr>
</tbody>
</table>

Study 2

<table>
<thead>
<tr>
<th></th>
<th>Rater 2</th>
<th>Rater 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rater 1</td>
<td>0.59</td>
<td>0.51</td>
</tr>
<tr>
<td>Rater 2</td>
<td></td>
<td>0.74</td>
</tr>
</tbody>
</table>
Table 3

Examples of essay questions used in Study 2

Definition
1. Define the concept of hemispheric specialization as it applies to the cerebral cortex.
2. Define reliability in the context of a psychological test.

Exemplification
1. Use an example involving Weber's Law to explain the concept of a difference threshold.
2. Use an example to show why a lie detector may not always detect lies.

Comparison and Contrast
1. Compare and contrast negative reinforcement and punishment.
2. Use comparison and contrast to present the important characteristics of the id and the ego.

Process Analysis
1. Explain the process by which the MMPI was constructed.
2. Describe the process of parent-child bonding.