LONGITUDINAL RESEARCH ON WRITING ASSESSMENT

Speakers: Paul Ammon, University of California, Berkeley
Catharine Keech, San Francisco State University
Introducer/Recorder: Juan Anino, San Francisco State University

Paul Ammon and Catharine Keech presented evidence which pointed to difficulties in the longitudinal assessment of writing development. Ammon began by stating that evaluators and researchers in the field, when engaging in longitudinal assessment, assume that students' writing "gets better and better," but he pointed out that writing sometimes "gets worse before it gets better," that "dips" often occur in student performance over a given period of time and instruction. These dips might indeed complicate the task of writing assessment.

Ammon discussed a study in which (over the period of a school year) the writing samples of bilingual children at the third and fifth grade levels were examined by means of both analytic criteria and holistic scoring methods. He found that overall scores generally did not show significant gains over the school year. The sum of the analytic scores was more sensitive to gains, but only among students who began the year at low levels of performance. In attempting to interpret these results, Ammon noted that some of the analytic scores actually went down, a factor which probably accounted for the general impression of little or no gain (as a gain on one dimension was cancelled out by a dip on a different dimension). Such dips could be explained by a number of factors: the student, when given a new version of the "same" task, may redefine it in a more complex way, newly acquired writing strategies may be overgeneralized to situations where earlier strategies would have been more appropriate in attempting something new; the student may experience cognitive overload which brings about deterioration in some aspects of his or her performance.

Ammon articulated some important implications from his research findings: 1) that pre- and post-test comparisons may be problematic, because even though "parallel" topics might be given, the tasks as defined by the writer might be different; 2) that multiple analytic scores are helpful but cannot simply be added up," because gains in one category might be cancelled out by lower scores in other categories, and 3) that, while we lack a theory of writing development sufficient to guide our use of analytic scores for assessment purposes, collection of longitudinal data may be helpful in forming such a theory. Ammon closed by offering a suggestion through which the testing model, with its assumption that writing "gets better and better," might be better aligned with the "teaching model," in which teachers should ideally acknowledge that students' writing often "gets worse before it gets better": In terms of assessment, students should be given credit for the positive changes which have occurred and should not be penalized for the negative changes which may occur at the same time.

Catharine Keech offered evidence of how these "dips" in scores appeared in a longitudinal assessment study of a high school population, and offered some explanations of why they occurred. During the six years of annual assessment at the high school, improvements in curriculum and approaches in writing instruction were being implemented. It would be natural to expect that through instruction and the natural effects of maturity, students' writing would get better, that a student's progression would "dip" from one year to the next for a given age group moving from freshman to senior status. But once again, assessment results showed dips in performance. Keech found that differences in readers' minds each year could not fully account for the counter-intuitive sequences which occurred in the longitudinal test scores of the students. A key element was how the students defined the writing tasks: what they attempted in response to the assignment. When the assignment or test question changed, were students able to draw on their knowledge in a variety of approaches? Or did they have access to only one approach which may have worked well for an earlier task but may have been less appropriate to a later assignment? Keech offered the analogy of finding one's way with or without a map. When asked to write narrative, students were already well-familiar with the terrain, had an internalized map, and could concentrate on the scenery without getting lost—for most, storytelling functioned as a "familiar well-marked freeway." Others, when they left the freeway, wrote something like oral dyad, not giving the reader a frame of reference or establishing for a reader why they are writing. To be able to "write an essay" successfully, the students need a new map, one through which they could move from a narrative sequence of events to a "point by point" form of discursive writing. Without a full understanding of this map and the ways in which it differs from other maps, the students may define the task according to the map they are familiar with, thereby producing narration or oral dyad, and receiving lower scores than on a previous task for which the old map was appropriate. Or, in attempting to "follow the map" closely, students may produce an "essay," but in the process forget that in producing the piece of writing they are supposed to "say something."

Thus, they may produce an essay which follows the format well but which is relatively devoid of insights or substance, and so may receive a lower score than an earlier, less formally-structured piece in which the student freely explored more interesting insights. In addition, the move into formal exposition and argument may require so much attention to map-reading—or figuring out the new discourse level rules—students may lose control over mechanics or fluency at the sentence level. Keech then presented a series of "personal reflection" topics through which students could progress from storytelling, to "tagging on commentary and reflection," to engaging in full analysis of a problem, in concept formation and extended abstract discourse. In concluding, Keech stressed that we need to build into the assessment process ways of recognizing and taking into account when the task has become more difficult—whether as a result of tester's demands or as a product of the student's own ambitions.

During the question and answer period, one audience member brought up the Dartmouth Study and asked how the findings of the two parallel studies related to that study. It was pointed out that as college students moved from role learning to abstract understanding, the quality of their writing went down, a finding which confirmed Keech's findings.

(Continued on page 22)
IMPACT OF TESTING (continued)

inherent conflicts between different kinds of test validity and of the competing purposes of the exam. The resulting tension between the course and the exam affected instructor-student relationships adversely, and these, in turn, resulted in demands from both instructors and students to change the exam. Initial objections to the exam focused on the validity of the measure. The instructors complained that the multiple-choice section of the original exam measured editing skills rather than writing and composing skills. In response to this argument and to the knowledge that research has been unable to demonstrate a relationship between awareness of formal grammar and improved writing skills, the program director eliminated the multiple-choice section and increased the amount of time students were given to write.

Students were allowed 35 rather than 20 minutes for the personal essay and 1 hour and 15 minutes rather than 45 minutes for the expository essay.

The instructors, however, continued to object to the exam, arguing that an impromptu writing task was an invalid measure of the writing skills that students developed in a process-oriented course emphasizing prewriting, planning, and revision. Students objected to the exam for another reason; the exam grade (1% of their final grade) often resulted in their receiving a lower final grade than they expected. Because of the deferred grading system, this grade affected two courses (8 units) rather than one.

Some prior research by Sanders and Littlefield and Rosemary Hake had indicated that an impromptu writing task underestimates a student's writing ability. In the spring of 1984, Michael Holzman, the previous USC Director of Freshman Writing, and Betty Bamberg conducted a study to compare the effect of giving students prior knowledge of the exam topic. They found that the average holistic score improved significantly when the students knew the topic and could discuss it before writing the essay. In the fall of 1984, the exam was changed to its present form. Bamberg has found that by gathering a small group of readings and announcing the exam topic ahead of time, the exam writers are able to develop more interesting and challenging questions. Not only are exam scores better, but also the responses are more interesting for graders to read. In addition, Bamberg believes the exam now more closely approximates out-of-class writing assignments and course objectives and, therefore, emphasizes the most important function of the exam: to provide a focus on course content and objectives.

In conclusion, Bamberg explained that the USC Freshman Writing Program will be restructured in 1986-1987 to eliminate the deferred grade and to create two separate courses. Bamberg hopes to keep the exam as part of USC's writing program because she believes its original purposes remain valid.

However, its form must be carefully designed to maximize the beneficial aspects and to minimize conflicting purposes.

IMPACT OF FIFTH GRADE TEST (continued)

and the third provides concrete aids for the writer. Examples of these prompts follow:

I. Topic—You have asked a friend to take care of a pet because you are going away for the weekend. Write a composition of about 150 words telling your friend about the pet and how to take care of it.

II. Questions—1. What kind of pet can you write about?

2. What are some things you want your friend to know about the pet?

3. What are some things that your friend should do to take care of the pet?

4. Think how you might end your composition.

III. Concrete aids for the writing

In preparing your composition, remember to:

1. Write your first draft on scrap paper.

2. Read carefully what you have written.

3. Make any changes that will improve your first draft.

4. Write your final copy on the answer paper given to you by your teacher. Use a pen to write your final copy.

5. Draw a line through any mistake you may make when you are writing your final copy. Make the correction and continue on with your final copy. You do not need to begin a new final copy.

6. Read your final copy before you hand it in to make sure you have not made any copying mistakes.

Mayher's and Lester's concluding comments focused on the impact of the test. The adoption of a statewide test, with the holistic mode of evaluation, has resulted in many benefits:

1. Statewide attention to the writing of elementary children by their teachers and administrators.

2. Reinforcement of the ideas that students must write early and regularly and that attention must be given to their development of a methodical approach to writing.

3. Introduction of in-service training in a number of elementary schools.

4. A new awareness on the part of teachers of the benefits of a coordinated attack within a school on the problem of improving student writing.

LONGITUDINAL RESEARCH (continued)

Keene discussed one of her studies in which the error counts stayed the same for pre- and post-test samples for a freshman composition group, but she noted that, in the post-test samples, there was a marked difference in the kinds of errors the students made. Another audience member pointed out that to do well on a proficiency test, a timed situation, students must have teachers who will "teach to the test." Keene affirmed this, noting the need for making students "test wise," and added that in the evaluation project she had helped lead, the students could not do well because within the course of their instruction, they had learned the importance of time as a necessary element in the writing process.