THE GREAT DEBATE: OBJECTIVE TESTS VERSUS WRITING SAMPLES

Moderator: Virginia Slaughter, Instructional Resource Center
Panelists: Kenneth Bruffee, Brooklyn College
Evans Alloway, Educational Testing Service

In her introduction, Virginia Slaughter emphasized the reality of the "Great Debate" by pointing out that three-quarters of all colleges and universities in the United States use objective tests "to some extent" to evaluate student writing, despite growing evidence that writing samples "have excellent reliability" and that most faculty prefer them. Then, by request of the panelists, participants broke into small groups to formulate questions they most wanted the panelists to comment upon. Some twenty questions were forthcoming; among them were:

- What are the pros and cons of objective versus writing sample tests, especially in regard to educational validity, fairness to students, practicalities of administration and scoring?
- What is the most reliable kind of objective test? Writing sample?
- What is the best method for administering and evaluating a writing sample? Can sample tests be designed to evaluate the process of writing as well as the outcome of writing?
- How can administrators and other faculty be convinced of the superiority of writing samples over objective tests?
- Do professional test-makers consider a school's curriculum when designing a test? Do they consult with faculty?

In his response, J. Evans Alloway made the following points: The best reliability comes from tests combining short answer with several writing samples; holistic grading of writing samples is reliable, the panelists provided that the evaluators are well-trained and able to apply agreed-upon criteria to their evaluative readings with reasonable consistency; writing samples have the advantage of being valid by definition, but short-answer tests must prove their validity; once proven, short-answer tests are more economically and efficiently administered and scored, thus making them attractive to administrators. The main thrust of Alloway's responses, however, was aimed at appraising participants of the roles that educators and those who influence educational policies play in the kinds of tests school systems select, design, or commission. The inference was that the "Great Debate" was not so much between those favoring objective, and those favoring writing sample testing, but rather a debate among the various constituencies that have a say in the policies of a particular school or school system. As Alloway put it: "The testing of writing is a social and political issue as well as an educational one."

Kenneth Bruffee expanded on the concept that the testing of writing is not a strictly educational matter. Referring to his experiences in helping to create the CUNY testing and evaluation program, he cautioned those involved in setting up such programs against the idea that they can operate without reference to the political, social, cultural, and economic forces that affect large-scale, mandatory testing programs. The testing program that a school finally decides on will represent a "big compromise" that reflects and, at least to an acceptable extent, satisfies the various "mandates" that are laid upon a school by interest groups exercising varying (Continued on page 8)
CREATING AND FIELD TESTING TOPICS
Facilitator: Richard Lloyd-Jones, University of Iowa

The problems in creating and field-testing topics were outlined in Richard Lloyd-Jones' opening remarks:

- Students may write badly because of poorly conceived topics.
- Too long a "stem" or stimulus question, may confuse students.
- We may harm students with poor tests and testing procedures.
- An essay test topic leads to a performance that is one datum (as opposed to fifty items on a multiple-choice test, of which one or more may be discarded).
- Sometimes within the constraints of a test situation, we cannot elicit from students the kind of writing we would like; a one-hour essay may not allow for profound thinking.

The participants then discussed these problems at length and suggested several possible solutions. First, faculty and administrators who design writing tests must avoid test bias (which occurs when the knowledge needed to do well on a certain test is more prevalent among certain student populations than among others).

On the other hand, in creating topics that will be appropriate for everyone, we must also try to avoid the banality of boringly general questions. Boring topics may result in greater discriminations among students because they frequently demand greater originality and cleverness. We should provide several interesting questions to elicit the best writing from all our students.

In addition, field-testing is very important for determining the demands of a test question. Field-testing a topic may be impractical if test security is an issue. But perhaps students should be allowed to know in advance the questions that they will write on; maybe they could even have a role in helping to formulate essay test topics.

Several participants also noted that test topics often emphasize the argumentative mode to produce an expository essay, but this tendency may limit students (and perhaps whole composition courses) to one rhetorical mode. Test questions requiring the use of other modes might be equally useful in evaluating writing.

Finally, the group suggested that the National Testing Network in Writing should provide a pool of questions to be used in testing and composition programs around the country.

Laraine Ferguson, Recorder
Bronx Community College

WRITING ASSESSMENT AND CURRICULUM
Facilitator: Barbara Weaver, Anderson College

Barbara Weaver introduced the session on writing assessment and curriculum by asking the audience to consider the following points:

- The relationship between assessment and curriculum
- Assessment within writing courses
- Assessment across disciplines

She also asked them to consider testing as an impetus for curricular change, given that a particular kind of testing can impinge negatively on curriculum. The group responded with questions about introducing a testing program and dealing with teacher anxiety, about training faculty readers; about the ways in which the assessment process has altered the traditional content of courses; about the issues of retention within remediation, appeals, repetition of courses and textbook choices; and about the restrictive implications of assessment tests on curriculum.

Weaver urged her audience to understand the politics at their respective institutions and to use friends within content area faculty to initiate groups of instructors interested in writing and the assessment of writing, so that readers might be found. She also pointed out that some kind of compensation should be paid to faculty who act as readers. In addition, she stated that the assessment process was bound to change traditional course content, one possible negative effect being teaching toward the test.

Weaver noted that on the other hand, schoolwide writing assessment increases the accuracy and the reliability of teachers' judgments of students' writing skills, and it keeps faculty and evaluators talking about the nature and definition of competency. It also communicates to students that competency is paramount. Students do far more writing, which in itself constitutes a profound curricular change. Assessment creates new tiers of remediation, thus having direct impact on curriculum.

There was agreement among audience members as to the need for a placement instrument; however, there was much debate about the need for exit testing. It was pointed out that testing tends to homogenize instruction and textbook selection. On this latter issue, there was consensus about the desirability of individual choices as long as there was a clearly defined syllabus.

Alfredo Villanueva, Recorder
Hostos Community College

GREAT DEBATE..., (continued)

degrees of authority and control over educational policies and practices. Bruffee advised those who take on responsibility for instituting writing assessment programs to start with the attitude that there are "absolutely no absolutes" that can be applied automatically to every writing assessment situation. Given the inevitable pressures from the various mandates that must be satisfied, educators must try to "manipulate the mandates" so that, despite compromises, the testing of writing "serves educational ends."

Both panelists made it clear that the "Great Debate" is not simply an "in-profession" disagreement among educators holding different views about the best way to test and evaluate writing ability. Both stressed that the best test for a school or school system is one that reliably assesses the writing ability of students while, at the same time, satisfies the diverse demands that arise from the larger community of which the school is a part. The question, "Why are we testing?" therefore, must be answered before the questions of "Which test?" and "Which evaluation procedures?" are addressed.

Domenick Caruso, Recorder
Kingsborough Community College