ESSAY TESTING AT CUNY: THE ROAD TAKEN

Why, if it is cheaper, neater, and politically easier to defend the use of an objective test to measure student writing, should a large, multi-campus university choose to assess student writing through the use of a writing sample? Administrators, misled by fad or ignorance, and the objective tests, seduced by the notion of "objectivity," and concerned with both cost-effectiveness and potential liabilities, used the wrong tests which can be read by computers. Why even consider the use of writing samples? Why, when faced in 1977 with the problem of mass testing of over 30,000 entering freshmen, did the City University of New York (CUNY) choose the less travelled and more hazardous road of the essay test as part of its Freshman Skills Assessment Program? 1
The answers to these questions began taking shape in 1978 soon after the Board of Higher Education mandated the assessment of basic skills. Fortunately, historical forces coalesced which enabled us to use an essay test as the sole instrument to assess minimum competency in writing. At that time, a dean at CUNY's Central Office had both CUNY-wide and national recognition for her research in basic writing. Mina Shaughnessy's seminal work in the analysis of student writing problems, her charisma, and her considerable influence within the University's central administration had an enormous impact on the direction of the testing program and its underlying assumption that testing is part of the larger issue of how best to teach our non-traditional students.
Mina Shaughnessy was also instrumental in identifying faculty from CUNY colleges to serve on the faculty task forces which developed the writing test. In Fall 1977, the Chancellor's Task Force on Writing was formed. Composed of nine experienced writing teachers and supported by those in testing and measurement, the Task Force was charged with developing or choosing an instrument to assess minimum competency in writing and with setting an appropriate level to determine that competency. Because of their considerable experience with Open Admissions and special program students at CUNY, the Task Force members, of which I was then a faculty representative, were able to agree upon certain basic assumptions and goals. We knew that testing inevitably also includes reading, and we felt strongly that the ultimate goal of educational assessment must be the improvement of teaching and learning. If there were going to be teaching to the test, we were determined to construct a test that would be desirable to write. We knew that we did not want to encourage thousands of students in writing classes throughout the University to sit in classes and fill in blanks in workbooks. We wanted a test which would signal both our colleagues in CUNY and our colleagues teaching English in the New York City public schools that the business of writing classes is writing.
We believed that a piece of student writing indicates real writing ability more accurately than a machine-scored test. We all had been hearing complaints from colleagues teaching English departments that our students couldn't write; pressed for a definition, they talked more about students' inability to stick to an idea, develop that idea, and illustrate it with specific examples than about the surface and mechanics of writing. An objective test, we felt, would measure and therefore encourage an educational focus on error correction rather than thinking and organizing skills.
Some members of the Task Force would have liked to give students even more than the fifty minutes which we allow them, but the essay test was to be part of a battery of tests which also included reading and mathematics, and testing time was a major problem. Some would have liked to ask students to write in several modes at several different times, but, again, logistics and time would not permit this. As is the case with all modes, we decided to ask students to write an expository essay. In the real world of writing there are few "pure" categories such as narration or description. Both are

The PROSPECTS AND PITFALLS OF A UNIVERSITY-WIDE TESTING PROGRAM

The issues of when, how, and whether to test writing skills have received a great deal of attention in recent years from legislators, school administrators, experts in testing and measurement, and, more belatedly, from teachers and students. Concern with these issues has generated an outpouring of essays and books—many of them impressionistic, a few of them informative—by which most of us have come to know a good deal more about testing. The debate is unlikely to slacken, since in the field of higher education alone, many states and colleges are still considering whether to implement inclusive testing programs in writing. However, an increasing number of institutions have decided to introduce new tests. Some universities, The City University of New York being the largest by far, have done so without assigning the work of test development or administration to a large outside agency; instead, they have placed these responsibilities on their own faculty. For these schools, a new question has emerged: how can we test writing and maintain it as a vital element in the educational life of the college? What are the pitfalls once a test has been introduced and officially accepted, and how can these pitfalls be avoided?

Colleges can anticipate problems for two reasons. First, any new policies affecting large numbers of people, unless the changes proved themselves to be utterly unworkable, assume an air of stability and permanence with surprising speed. This is especially true when, as in the case of testing, policy is tied to the calendar; its impact is felt at predictable times and it comes to share some of the inevitability of the seasons. The CUNY Writing Assessment Test (WAT), a matter of considerable controversy within the University before it was introduced, has already, in four years time, settled in as part of University life. The test is given at stated times on each campus, read at stated times, and the results are officially audited each June. Teachers who opposed it may now find themselves pleasantly surprised by its usefulness or resigned to it as an inconvenience, but their thoughts and their energies have turned to other issues on their campuses. The WAT now seems as solid and immovable as the filing cabinets in which the test results are stored. Furthermore, one of the crucial elements in testing is reliability, the assurance that students taking or retaking a test will find it consistent in measuring what it measures. That is not the same as saying the student will be taking

(Continued on page 16)
an identical test but it does exert a strongly conservative influence on a testing program. The idea, sensibly enough, is to replicate as many of the features of the test and of test-taking as possible. In doing that, test makers can easily find themselves tinkering with surface features and nothing more, preparing Form B, Form C, and so on through the alphabet, without ever thinking again of more far-reaching alternatives.

The major pitfall, then, for a testing program once the test is in place is evoked in the very words "in place"—the risks of stasis, of habit, of benign neglect. One has only to participate in a reading where training procedures are scented or ignored because "the readers already know all about that" to realize the consequences of familiarity breeding inertia, if not contempt. The question is how can a testing program be continually re-energized without being thrown into confusion?

A successful testing program first requires a change of attitude on the part of those who sponsored the program and now direct it. During the phase of debate and initial implementation, the proponents of the new program may find themselves shouting louder than anyone else—that is, they may succeed by insisting vehemently on the superiority of their proposals. Or they may find themselves speaking more softly than anyone else—that is, they may succeed by claiming most conflicts to be inconsequential or based on misunderstanding. But however they proceed, they invariably welcome the quiet that falls once their program has become part of the way things are. Sponsors of testing programs have to recognize the dangers in this silence of acquiescence and adjust their methods accordingly. They must stimulate new discussion and keep the testing program alive by fostering some degree of renewed debate as part of a search for improvement.

There are various venues in which colleges can assure the continuing vitality of their testing programs. Some have been demonstrated at The City University, which I use for illustration because it is the institution with which I am most familiar. First of all, CUNY has rotated faculty appointments on its policy-making boards for testing so that new members join the boards each year, able to look at the Writing Assessment Test from fresh perspectives. That practice, combined with CUNY's willingness to turn over most aspects of the testing program since its inception to faculty control, has provided a strong base for continuing faculty involvement in the test. Furthermore, the writing faculty who designed the WAT subsequently developed a writing curriculum to serve as an example of how testing might support the needs of college writing programs. This sample curriculum continues to generate interest at individual colleges in alternative ways of adapting the WAT to classroom use. Faculty can also be involved more meaningfully in a testing program when ways are found to keep testing and related matters a part of the faculty's academic interests. CUNY again provides some examples, for it has made data and writing samples from the test available to teachers who want to do serious research on various features of student writing and of test content. CUNY has also helped to fund such research and has offered seminars for teachers who are interested in learning more about basic research design and methodology. When an institution acts in this way, it confers a new importance on matters that previously carried little academic prestige for English teachers. At the same time, it prevents the test, its content, and its evaluation from becoming a bureaucratic preserve, another admini-

Robert B. Lyons is Associate Professor of English at Queens College, CUNY.

NOTES from the NATIONAL TESTING NETWORK IN WRITING

The City University of New York and the Fund for the Improvement of Postsecondary Education

Directors, NTNW
Karen Greenberg
Harvey Wiener
Richard Donovan

Editors, NOTES
Donald A. McCuade
Virginia B. Slaughter

Published by the Instructional Resource Center, CUNY Office of Academic Affairs. Marie Jean Lederman, Director