
TEST ADMINISTRATION, COSTS, AND DATA ANALYSIS

Speakers: Michael Ribaldo, City University of New York
Anthony D. Lutkus, New Jersey Department of Education
Introducer/Recorder: Frank Hubbard, Cleveland State University, Ohio

Michael Ribaldo, who directed CUNY's Freshman Skills Assessment Program, began the session by explaining how CUNY's testing program grew and discussing problems encountered during its development. CUNY is the nation's third largest system, with twenty-one campuses (seven two-year colleges and ten four-year colleges, plus a graduate school, a law school, and two medical schools). There are 180,000 students, 10,000 faculty, and 10,000 support staff of all kinds. The whole is governed by a board of fifteen trustees, divided between gubernatorial and mayoral appointees. In 1976, the trustees passed a resolution to assess the abilities of

students at the point of transition between the sophomore and junior years (sixty-one credits). 1976 was also a time of financial problems for the system: 150 faculty lost their jobs and everyone suffered a two-week furlough without pay. The resolution appeared to some as a way to cut costs by shrinking the size of the university. CUNY's Chancellor Joseph Murphy responded to the resolution by appointing a faculty committee to advise him. The committee supported the assessment of reading, writing, and mathematics skills, but surprisingly also suggested that such testing be shifted to the start of academic careers, when it could be used for placement and diagnosis, and then be repeated prior to the junior year to test for competency. The Chancellor appointed three task forces, one in each area. The reading task force decided on a commercial test, using twelfth grade scores as cut-off points; the mathematics task force decided to construct its own tests, as did the writing task force. The writing task force modified the ETS advanced placement model and developed its own grading scale. The writing test is a holistically scored fifty-minute essay. The pass/fail cut-off was set for the university as a whole, with individual campuses allowed to raise it if desired for local placement

decisions. Students who failed had to take basic skills courses in their respective colleges.

The biggest problem facing the testing program has been its cost. Since 1978, every student entering CUNY has taken all three tests; 33,000 take the tests each fall, and an additional 10-12,000 are tested each spring. Moreover, these tests are used as exit tests from the basic skills sequences and also given to transfer students, raising total administrative costs considerably. In the 1976 retrenchment, grading was made part of the faculty's workload. Campus presidents felt some hostility from their faculties, however, so a patchwork of compensation alternatives developed—released time, exemption from registration duties, pay as adjuncts, or no compensation at all—which remains in effect today. One college resisted, and the faculty union took the matter up with the central office; the faculty lost, but the patchwork arrangements had already been made elsewhere and the precedent was therefore never enforced.

To insure uniform standards for scoring the writing tests, a yearly audit is conducted. The most reliable readers in the university rescore 100 freshman tests and 100 retests from each campus; these 3,400 tests get read every June by 70 faculty using standard norming with range-finding papers. Meetings are scheduled at colleges to address errors of either severity or leniency. The audits have generally been satisfactory: The agreement on pass/fail marks averages around 83 percent. Two readers agree or are only a point apart on 90 percent of the exams. In addition, several CUNY faculty have conducted studies of the validity and reliability of the writing test, also with satisfactory results. Research on the test's impact on course structure and curricula is under way.

The costs of assessing writing at CUNY are hard to estimate because of the variety of compensation practices, but in readings where the compensation is at the adjunct rate, \$20 per hour, readers generally read twenty exams per hour. Each reading, then, costs \$1, and because each paper is read twice, and 23 percent are read a third time, the total is an estimated \$2.23 per test. Another reason costs are hard to estimate with precision is that the 60 percent of exams that are administered in the spring are relatively cheap to score compared with those in the smaller summer administrations. During the school year, topic development and field-testing are performed by faculty committees, and a writing task force develops prompts, pretests them, and analyzes the results without costing real dollars. The chief reader for each test scoring session is usually a deputy chair of an English department, who handles the selection of readers, their training, the mechanics of scoring, and the transmission of scores and papers for audit to the central office.

CUNY does little central data analysis except on initial test results. To avoid over-centralization and because data processing is too costly, individual colleges keep track of retests. University data indicate, however, that freshman writing scores have declined since 1978; now only 51 percent are able to pass upon entry to the university. It was hoped that the Regents examination required for high school graduation would help resolve the basic-skills problem, but that test has apparently not been enough. The attrition rate at CUNY, as elsewhere, remains high, but early exposure to basic skills instruction has been shown to be one way to keep students in college beyond the first year. Interestingly, Ribaudó noted, among those who do remain and want to continue their studies, no lawsuit has been filed for non-admission to

the junior year. Those denied are few and, generally, are no longer on campus. Among those who take the junior year entrance exam, the best predictor of success is generally the ability to pass the basic skills courses the first time around.

Tony Lutkus, the director of New Jersey's statewide Basic Skills Assessment program, noted that New Jersey is much like New York in many ways: New Jersey admits about 50,000 students per year to thirty public and twenty-six private colleges. The chronology of the basic-skills assessment is also similar: a trustees' resolution in March 1977 responded to faculty discontent with the preparation of students and concern over open admissions. The public colleges were asked to develop a test and to offer remediation. The results of the tests were to be made public (and will soon be made available to boards of education across the state). Colleges were to place students, remediate them, and follow up on their academic careers. The testing package includes a twenty-minute essay graded holistically on a six-point scale, a fifty-minute reading comprehension test with fifty questions and seven pretest items, a thirty-five-minute test of sentence sense, and a forty-minute computation (fifth grade arithmetic) and elementary algebra test (ninth grade). As at CUNY, the test is given after admission, but before registration.

On the sentence-sense test, the aim was not a normal distribution of scores, but rather the left wing of such a bell curve; in other words, the test was supposed to be easy for well-prepared students, who should pass it in great numbers. The holistically scored writing test, unfortunately, seems to result more closely in a bell curve. Placement in writing classes uses multiple indicators; the central office simply sets minima (e.g., a seven on the essay). The results, "homogenized" by ETS, are reported in a brochure. The brochure for 1985, for example, reports that 34 percent failed the verbal proficiency sections, 40 percent failed some part of the verbal proficiency assessment, and only 26 percent were fully proficient. These numbers are contrasted with the high school graduation numbers: 29 percent failing all, 43 percent failing some, and 28 percent testing as proficient. This comparison suggests that the high schools are the location of the difficulties. How, then, can the college system help? Administrators are providing assistance in the form of workshops and seminars, and they have also included high school teachers in the grading process. New Jersey reads its 50,000 essays twice, but under conditions that contrast markedly with CUNY's. Gathering at a luxurious retreat in Princeton, 70 graders, paid \$75 per day plus travel and room and board, grade 15,000 exams. ETS supplies the data analysis and statistical consulting. This year, the cost was at least \$177,000, or \$3.10 per student just for the essay section. One of the justifications for this expense has been the cooperation between high school and college faculties, and the resultant good will, according to Lutkus, has seemed cheaply bought.

Much research is under way to assess New Jersey's test essay. Among studies of the impact of variable factors—topics, students, test sites—on test results, one finding shows that as much as half a point depends on the topic. To check this, an old topic is being reused this year, and some old papers are being regraded. Another project is seeking to determine whether a thirty-minute writing period would be better, and is giving some students the usual twenty minutes, others thirty minutes but with ten given only to planning, and still others thirty minutes

with no restrictions. Paralleling site-variation studies in Washington on the General Education Diploma (GED), New Jersey is also assessing reliability between sites. The academic progress of students with second-language backgrounds, and of students who fail the test, is also being studied.

In one study, New Jersey's state colleges defined a cohort of students and tracked those who needed no remediation, those who completed their remediation, and those who did not complete it during the course of four semesters. Factors considered were pass rates, retention, the relation of credits attempted to those earned, grade point averages, and results on pre- and post-tests in the remedial courses. After four semesters, 70 percent of non-remedial students, 75 percent of remediation-completed students, and 39 percent of the incompletely remediated were found to be still in attendance at four-year state colleges; the numbers at two-year campuses were 51, 55, and 22 percent, respectively. In terms of numbers of full-time students, over the two school years from fall of 1982 to spring of 1984, this cohort began with 18,000 regular, that is, non-remedial, students and ended with 11,872 (10,086 of whom had C averages or better); 5,200 students assigned to remediation completed it, and

3,655 of these were still enrolled in 1984 (2,253 had C averages or better). There were 3,016 assigned to remediation who did not complete it, and only 871 of these were still around two years later (500 had averages of C or better). The conclusion is inescapable that remediation is crucial to the academic survival of this student population.

This conclusion has been reinforced by findings of the New Jersey Department of Education, which has established that, while SAT scores nationwide have risen recently, no comparable rise has taken place in the population which takes only the New Jersey test. The responsibility for remediation, however, rests with individual colleges. According to Lutkus, the successes of the New Jersey Basic Skills Assessment program include problem identification, increased services to students, an extensive data base, raised standards, upgraded program evaluation, and improved communication between secondary and post-secondary educators in the state. But the state department does not control the content of remedial courses, though it has had to address the issue of degree credit for such courses and has reduced the number of such credit-granting campuses to three. ■