VALIDITY ISSUES IN WRITING ASSESSMENT

Speakers: Sybil Carlson, Educational Testing Service, New Jersey
Barbara Gonzales, CUNY Office of Academic Affairs

Introducer/Recorder: Carolyn Leitman, Cuyahoga Community College, Ohio

Sybil Carlson described an ongoing research project investigating the validity of the Analytical Reasoning section of the Graduate Record Exam (GRE). The project, which investigates the extent to which the assessment of reasoning skills is confounded by the assessment of writing skills and the extent to which they can be differentiated, is motivated by the need to demonstrate "construct validity." Carlson said that many tests do not have construct validity: there is no evidence that many tests measure what they were intended to measure.

Writing samples were selected for observation because the thinking or reasoning skills of post-secondary students are largely evaluated by how well they can express themselves in written form. The data, consisting of writing samples generated by beginning graduate-level students, are being examined to identify reasoning skills that are relatively independent of other skills. Carlson described some of the relationships among scores assigned to samples according to the various scoring schemes that were developed. Essays were evaluated for holistic writing scores, holistic reasoning scores, two different reasoning scheme scores, Writer's Workbench scores, and subscores on the GRE General Test of Reasoning.

Carlson indicated that this project enabled researchers to observe consistent patterns of relationships among different assessment instruments, and also to find out what the different schemes developed to assess reasoning skills actually can and cannot do. This research may eventually lead to the development of useful measures for evaluating reasoning skills, measures that can also provide students with valuable feedback on improving their writing/reasoning skills.

Barbara Gonzales discussed the assessment of the writing skills of ESL students at CUNY. Gonzales began her presentation with a general overview of the ESL population and programs at CUNY. In 1984, 14,000 students enrolled in CUNY's various ESL programs. All students (native speakers and ESL students) must pass three basic skills tests in order to enter the junior year of college. The writing section of the test consists of an essay in which the student must agree or disagree with a given statement. Students must compose and revise their responses within fifty minutes, the equivalent of a single class period.

Gonzales explained the rationale for asking foreign students to submit a writing sample. She pointed out that elements of good writing, such as clarity, organization, correct use of idiomatic English, proper word choice, etc., would not be revealed by testing students with a multiple-choice test. Moreover, the test serves the overall goal of the ESL programs--to achieve communicative competence in writing--as well as that of writing teachers, who define competent writing as the coherent expression of thoughts. However, CUNY policy requires that the writing tests of ESL students be read by at least one reader familiar with the writing of native speakers of English. This guarantees appropriate and correct recognition of errors pertinent to these writers.

Gonzales also discussed the relationship between the testing and the teaching of ESL students at CUNY. The CUNY Task Force on Writing recently developed a curriculum model for ESL in order to incorporate these measures. All students for college-level work. Writing proficiency, as defined by CUNY's test, was designed as the ultimate goal of the ESL program. Instruction proceeds through four levels, and the importance of reading academic material is emphasized beginning in the very first course. At all seventeen participating CUNY colleges, ESL teachers stress the importance teaching functional communicative competence in writing, and they are collaborating with Basic Writing teachers to develop standards for writing proficiency for all students.

TEST ADMINISTRATION, COSTS, AND DATA ANALYSIS

Speakers: Michael Ribaudo, City University of New York
Anthony D. Lukens, New Jersey Department of Education

Introducer/Recorder: Frank Huband, Cleveland State University, Ohio

Michael Ribaudo, 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 189,000 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 10,120 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 reorganization, grading was made part of the faculty’s workload. Campus presidents felt some hostility from their faculties, who resented the increased administrative burden. The faculty, in turn, resented the lack of adequate compensation 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 position that 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 norms 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 85 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, $25, about 12,500 readers, each reading about 20 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 $2.25 per test. Another estimate is that only 60 percent of the 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 scores. 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 1970s; now only 23 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, Ribaudo 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 Lukus, 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. In March 1977 respondents to faculty disconsolate 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 test (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 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 itself, however, 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 26 percent testing as proficient. This comparison suggests that the high schools are the location of 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 59,000 essays twice, but under conditions that contrast markedly with CUNY’s. Gathering at a luxury resort 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 Lukus, 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—and 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
IMPLEMENTING WRITING COURSES AND TESTING STUDENTS FOR MINIMUM COMPETENCIES ACROSS THE CURRICULUM

Speakers: Lana Silverthorne, University of South Alabama
Marc Maire, University of South Alabama

Lana Silverthorne opened the session by explaining the development of the Writing Across the Curriculum program at USA, an institution of approximately 10,000 students. The program began in 1979 with the appointment of a twelve-member Committee on Writing. A survey of faculty members by this committee indicated a campus-wide awareness of writing across the curriculum programs; thus, a University Writing Program Advisory Committee was appointed to initiate several changes at USA. The first of these changes was in the basic freshman grammar course (101), which was restructured into a writing workshop, using student papers as the text, with sequenced writing assignments designed to move students from informal, exploratory writing to more formal writing. Students in this course photocopy their papers to share with members of small groups, and they discuss the strengths and weaknesses of their writing. Students with more serious writing problems are also referred to the Writing Lab, though faculty members are discouraged from requiring Writing Lab visits.

Students are placed in 101 with an ACT [American College Testing Service] score of twenty-two or lower, which exempts only about 8 percent of USA’s freshmen. This course is given four hours of credit, but it is taken pass/fail so that students may repeat it with no penalty to their average. To pass, students must receive a satisfactory rating on a holistically scored exit exam at the end of the quarter; approximately 70 percent pass the first time. Passing enables them to go on to the next required writing course (102), a more traditional composition course for which students receive not only credit but a grade.

In addition to changing the freshman writing course, the committee developed two pilot programs. One is a university writing curriculum and the other a writing placement test for transfer students. For the writing curriculum, each undergraduate department selected one full-time faculty member who was teaching a course required for a major or minor to develop a writing requirement for her or his course. Faculty members who agreed to add this “designated writing credit” (W) to their courses were paid a one-time stipend of $400 and were guaranteed both reduced teaching loads and reduced student loads. In return, they agreed to attend a week-long workshop and incorporate the ideas of that workshop in their syllabi. Using the ideas implemented in the 101 course, workshop participants were guided in preparing sequenced writing assignments for their courses. These assignments encouraged frequent, short writings rather than the more traditional essay test or the one-time term paper. In addition, workshop participants spent time holistically scoring sample student papers to develop their skill in evaluating student writing. Forty faculty members attended the first workshop, and since then, nearly eighty more have participated in the annual workshop. As a result, 138 junior and senior courses carry W credit, with approximately sixty such courses available each quarter. Since the 1983-84 school year, students have been required to take two of these upper level W courses, at least one of which must be in their major or minor.

Marc Maire described the second pilot program, for which he had just completed a computerized study. The survey of faculty members had also identified a “transfer student problem,” that is, faculty members felt that many transfer students were poor writers. Consequently, a placement test was developed for any transfer student with the equivalent of USA’s 101 course from another school. This exam is much like the exit exam for the 101 freshman course, since passing this test means the same thing as passing English 101. For this placement test,