MODELS FOR THE EVALUATION OF REMEDIAL WRITING PROGRAMS

Speakers: Walter Cmielewski, County College of Morris, N.J.
Anthony Evangelisto, Trenton State College, N.J.

Introducer/Recorder: Robert Cirasa, Kean College of New Jersey

Walter Cmielewski began the session by presenting a theoretical context for program assessment. Quoting a 1978 study done for the National Center for Higher Education Systems by Leonard Romney, he referred to the still primitive state of educational assessment in general, to the important difference between the quality of program activities and the measurement of program outcomes, and to the very real possibility that program impact cannot be meaningfully assessed at all. At the very least, he noted, assessment is a complex, perhaps unmanageable task.

He then pointed out that people involved in writing program assessment should start by distinguishing between a pair of popularly synonymous terms: "remedial" and "developmental". Remedial education denotes efforts to correct specific behavioral weaknesses (i.e., writing skills alone). Developmental education emphasizes the whole person instead and attempts to develop strengths as well as to address weaknesses. All remedial programs ought to be developmental as well, but all developmental are not necessarily remedial programs. Moreover, Cmielewski urged program evaluators to consider whether they want to conduct summative evaluations or formative evaluations or both types.

A summative evaluation involves the measurement of observable outcomes; the emphasis is usually upon identifying any defective results of the program. A formative evaluation assesses what happened and how results were achieved; the emphasis is upon improving the program and keeping it moving effectively and efficiently towards its goals. Cmielewski added that formative evaluation should assess program efficiency as well as effectiveness, and that caution was that often programs lack clear goals and objectives and data is collected without regard for efficiently managing and reporting that data.

Having established a broad theoretical context for program evaluation, Cmielewski yielded the floor to Anthony Evangelisto, who observed that theoretical models for program evaluation have usually followed—not preceded—actual program development and that there are no generally applicable models for program evaluation. Individual programs usually evolve their own models, and they start by asking themselves some basic questions: "How good is this program? What makes it good? How can we make it better? What's wrong here? So what? Who cares?" Eventually, two basic concerns arise: What was the program supposed to accomplish and how successful was it in achieving its goals? To answer the first question, one must look at the specific goals of the program and at how clearly they are defined. To answer the second, one must gather data and interpret them.

There are three basic features that can be examined: the program (organization, materials, personnel, curriculum, facilities, etc.); the processes (methods, interactions, interventions, etc.); and the products (cognitive and affective learning, skill development, self-concept, motivation, etc.). Some of these products are very difficult to measure, but complete program evaluation requires that one resist the natural inclination to limit inquiry and answers to only those things that are easily quantifiable. Evangelisto then went on to discuss three models of evaluation and their flaws.

Type I is represented by the noted WPA Model [WPA: Writing Program Administration, 4 (Winter, 1980)]. It uses a battery of specific questions for the self-study of curriculum, faculty development, and support services. Reassuming the approach of regional accrediting agencies to overall institutional assessment, this approach focuses primarily upon program matters.

Type II is represented by a number of locally developed models that seek only to document gains in student skills. The Individualized Language Arts Program of Woahawk, N.J. for instance, uses holistic scoring of writing samples to study program effect, the New Jersey Writing Project of Monmouth Junction and South Brunswick Township, N.J. do likewise. In the same vein, the Hillsborough County, Developmental Writing Program in Florida assesses improved writing skills in terms of higher achievement scores in reading comprehension, language mechanics, and language expression. And the Elementary Writing Program of Fairfax County, Virginia, bases its evaluations on data from questionnaires, interviews, and a case study of an individual school. Evangelisto noted that these and other programs of this sort are routinely listed in the latest edition of Educational Programs That Work. In general, they all focus upon product, although there is some attention paid to program matters and processes as well.

Type III is represented by the work of the New Jersey Basic Skills Assessment Committee, which has recently begun to ascertain the effectiveness of programs through a combination of data collection and site visitation. Since 1978, a mandatory annual report, designed by the Assessment Committee, has measured data from programs in every public institution in the State on the following indicators of program effectiveness: (1) number of students tested for placement; (2) placement criteria; (3) number of students identified for remediation; (4) levels of remediation; (5) enrollments in remedial courses; (6) passing rates for final level remediation; (7) percent passing class who also perform acceptably on a single standardized post-test (usually a holistic essay); (8) differences in retention among those who needed no remediation, those who did but completed it, and those who have not completed it; (9) differences in GPA's among the same groups; (10) differences in total credits earned among the same groups; (11) differences in percent passing first college-level writing course among the same groups; (12) differences in "successful survival rate" (percent of originally tested students remaining two full years and performing with a cumulative GPA of 2.0 or better). Recently added to these quantifiable measures have been routine site visitations that seek to identify the local features of programs that are responsible for varying degrees
of effectiveness. Specifically, the visitations characterized: (1) program administration; (2) policies and procedures; (3) remedial curriculum; (4) remedial pedagogy; (5) faculty development; (6) instructional support; (7) student support; and (8) local program evaluations. The overall emphasis of this approach is equally upon product, process, and program, and through its multiple perspectives, it comes close to giving an accurate assessment of writing programs.  

**ASSESSING WRITING PROFICIENCY IN THE PROFESSIONS**

Speaker: William Lutz, Rutgers University, Camden, N.J.

Introducer/Recorder: Louise Silverman, Ocean County College, N.J.

William Lutz, who is a lawyer as well as an English professor, emphasized that law students usually do only minimal amounts of legal writing while they are in law school. Because of the high student-teacher ratio in law schools, law students are only required to write a few papers, final exams, and moot court briefs.

The audience for legal writing tends to be the legal community. Although, in theory, lawyers write for clients or for the general public, in reality, they write for an audience of other lawyers or for judges, who expect legal jargon in legal documents. Use of legal jargon gives access to the "sacred priesthood" of lawyers, but the use of jargon may prevent clients from understanding the content of a legal document. In addition to jargon, legal prose often contains multiple negatives, imbeddings, and prepositional phrases, as well as passive construction, Latin or Middle English phrases, and formal, ritualistic phrases. While special reading strategies are needed for reading legal documents, obscure, ambiguous prose is potentially misleading and harmful.

According to Lutz, legal writing should be the writing used in everyday affairs, with the usual attention to audience, purpose, and context. Law students' writing, therefore, should be assessed by lawyers for form and content. Simplicity and clarity should be the primary concerns in legal writing; a classic example of such simplicity and clarity is Justice Brandeis' dictum: "Danger cries out for rescue." While primary trait scoring can be used for legal writing assessment, Lutz prefers holistic scoring or portfolio assessment. Students can be trained through the reading of models of clear, legal prose; however, topics must test legal knowledge as well as writing skills.

**EXPLORING WRITING ASSESSMENT IN ANOTHER CULTURE: CHINA AND AMERICA**

Speaker: Marie Jean Lederman, LaGuardia Community College, CUNY

This session's premise was that the values of a culture are revealed through its examination system or systems. As she discussed writing assessment in both ancient and contemporary China, Marie Jean Lederman invited participants to examine similar issues in their own testing programs.

Lederman began by describing China's ancient Imperial Examination System, which lasted some 1,300 years. This first attempt to create a merit system for government positions was the earliest writing assessment program in the world (Sui Dynasty, 589 A.D.-618 A.D.). Candidates wrote essays explicating the writings of Confucius, a body of material which set forth the moral and ethical basis of society. These examinations were given on three levels, culminating in the capital examination in the Forbidden City in Beijing for the top three hundred candidates.

Among the issues which examiners faced was reader reliability, and so the procedure of using two independent readers and a third in case of disagreement was developed. Other issues were the problem of the relationship between the fixed form of the examinations and the creativity of the examiners, rest test policies, the control of cheating, and equity. These problems are reflected in literature widely separated in both time and space: the novel The Scholars, written by Ching-Tsu Wu around 1750 and John Hersey's The Call.

Lederman briefly traced the imperial examinations through 1905, when they were finally abolished in China. She then focused on the period immediately following the Cultural Revolution when the National Unified Entrance Examination for Institutions of Higher Education began (1977). The last years of the Cultural Revolution were a time of great educational ferment, centering on college admissions standards and entrance examinations. What types of assessments should be used? Who would be included and excluded? Who should be in the universities? This controversy was also reflected in the popular media, evidenced in the emotionally-charged Chinese film called Breaking with Old Ideas (1975).

The question was resolved in favor of the academically demanding university entrance examination which has become more stringent every year since 1977. It includes a section on Chinese language and literature—half of which is an essay. Lederman distributed a Review Syllabus for the Chinese language and literature section, and participants noted similarities and differences between the demands of writing assessments in China and the United States. The discussion focused not only on criteria for writing but on what these criteria imply about the values in both cultures.