POLICY AND ASSESSMENT

Facilitator: John Alexander, *Ferris State College*

John Alexander opened the session on Policy and Testing by stating that departments should clarify broad curricular objectives in order to justify establishing a comprehensive testing program. In addition, faculty should be aware that the benefits of a testing program go far beyond accuracy when a writing sample score is merged with an "objective" test score, such as the ACT English. Group training sessions, essential for the success of any writing sample assessment program, also have the effect of achieving some common understandings about evaluation and curricula, thereby improving instruction in large multi-sectioned courses. A message is sent to students and faculty in secondary schools that effective writing rather than correct grammar and syntax should become the measure for success at the college level.

Yet, for all the advantages that may result from testing with writing samples, there still is no evidence that the test adds to our ability to predict accurately student success in college. After a two-year study at Ferris State College using a computer-based monitoring system for writing sample administration, there emerges little evidence that writing sample scores can predict student success more accurately than standardized ACT scores. Statistical studies generated from criterion-referenced analysis of the writing samples administered to 1,200 incoming freshmen reveal that standardized tests and writing samples correlate well with each other and equally well with student grades through a writing sequence. However, discrepancies at the higher end of the scale of student ability as measured by the ACT suggest that writing samples are necessary for exemption and proper placement of high-ability students.

Furthermore, the Ferris State study showed that numerous problems endemic to writing sample assessment can be remedied. A computer-based monitoring system can equate student writing performance with national test norms, can adjust reader scores when marked divergence from rating norms is detected, and can monitor levels of rater agreement and estimated reliabilities. In addition, departments can obtain invaluable assistance from testing offices by running follow-up studies of student performance. Such studies can help to allay the age-old myths about testing and student performance that often get in the way of inter-departmental cooperation. They can also demonstrate that English faculty do not have an inherent distrust of technological data if used to serve humanistic goals.

Alexander then questioned whether we have, as a profession, really progressed that far by moving from machine-scored objective tests to writing samples? Clearly, English faculty have made their voices heard on and off the campus. But how far have we come in substituting for tests of usage and correctness tests that assess written products rather than processes? How much weight should we give to an instrument that measures only a single facet of an enormously complex set of skills?

Alexander concluded by saying that policy statements on campuses should reflect the following considerations:

- Who shall evaluate?

- How can colleges and universities minimize the problems inherent in the several methods currently used in testing?

- Should English departments do it on their own? How close should composition specialists work with measurement people?

- What complications arise when large multi-purpose campuses are asked to adopt uniform testing procedures monitored by a central administration?

Lila B. Soll, *Recorder*
Baruch College