TRAINING READERS
Facilitator: Robert Christopher, Ramapo College

In his presentation, Robert Christopher maintained that the selection, training, and support of the reader are vital to the quality of assessment and to the curriculum. Since the reader acts as a mediator between the institution, the test itself, and the student, only by emphasizing reader training can the integrity of the process and the interests of the student be protected.

Assessment reading is contextual, and there are few absolutes. It is, however, clear that the process will differ somewhat according to the leadership of the institution and the mission of assessment. Nevertheless, there are shared aspects of the experience common, for instance, to the CUNY collegial: the scale employed, the role of the reader, and the training and selection of readers. Although reading can be on a mass scale, an institutional scale, or a departmental scale, the constant throughout is "collaboration to reach a consensus through comparative judgments."

The reader has a dual responsibility: to validate the test and to interpret fairly the student's performance on it. He must assign a score within the limits established by the test, and that decision will determine the student's placement within the curriculum. In consequence, poorly trained readers may misinform students as to their performance, invalidate a test, and undercut the placement process.

It is clear that good readers are necessary if assessment reading is to have any meaning, but what are we to look for in our model of a good reader/reader? Reading holistically requires, indeed demands, a combination of skills and attitudes: reading acuity, a willingness to read supportively, the ability to use the entire range of the evaluation scale, skill in the use of available time, and a commitment to working collectively and collaboratively. Furthermore, a reader must be consistent with the institution for which he is doing the assessment reading.

Dr. Christopher suggested that institutions form a "pool" of readers which would work in solitude, sharing judgments and developing mutual trust. Such a pair would constitute a "dyad," and six readers, or three dyads, would form a core. This core of six should undergo intensive training. If necessary, the members of the core group would become "brokers" to bring others into the process. They could be paired off in new dyads so that the circle would be expanded. In order for this training to be effective, the core group would have to be "complete insiders," aware of how the institution functions, why and how topics are chosen, the method of scoring and grading, the way in which the scoring process is conducted, and the consequences of assessment.

Although it is clear that reader training is vital to the assessment process, it is hard to be absolute or specific about the nature of that training. Much depends, for example, on whether the scoring mode is holistic or analytical. However, Dr. Christopher emphasized the importance of a commitment to group work and to the process of trial and error. Naturally, readers who work within an institution that shows its appreciation of their efforts by providing a comfortable working environment and adequate compensation are more likely to maintain high standards of performance. But beyond these purely material considerations, an awareness of the purposes of a particular test, and, indeed, of the purpose of the entire testing process and the vital role of the reader within it, is

HOW TO BEGIN A TESTING PROGRAM
Facilitator: Robert Lyons, Queens College

Robert Lyons enumerated five questions to consider when thinking about a testing program: what circumstances produced the need for a test; what groups should be involved in creating the test; what is necessary to get them involved; how important is what; what is a reasonable timetable; and finally, what kind of outside help should be sought and from whom. He then opened the discussion to the audience, and what follows is a summary of that discussion.

The CUNY/west Assessment Test (WAT), was related to the advent of Open Admissions. Because graduates of two year units have the right to transfer to four year units, there was pressure for a uniform minimal competency exam that students must pass by accumulating 60 credits. One effect of the CUNY test was to strengthen remediation programs in all the units. Another effect of the WAT (which requires students to write an essay) was that it made colleagues in other departments more aware of writing and willing to participate in writing across the curriculum projects. Minimal competency tests should not be confused with proficiency tests. They are not meant to suggest that further writing instruction is unnecessary.

Several factors were involved in CUNY’s decision to use an essay test. The CUNY writing test was vehemently opposed to objective tests and through organizations like the CUNY Association of Writing Supervisors (CAWS) spent several years convincing the administration of the need for a writing sample. Many of the CUNY units had developed placement exams that required writing samples so that a tradition already existed. CUNY’s intention was also to influence writing programs of the New York City high schools.

It was stated by a member of a team that creates exams for admission to medical schools (MCAT) that a writing sample is added in order to send the message to aspiring medical students that medical schools also care about writing.

A major goal in creating a test is to avoid having too many demands placed on it to keep its major purpose in mind. A compromise must be reached between a proliferation of tests and using a single test for purposes for which it may not be suitable. At CUNY the WAT is often used as a placement exam as well as an exit exam. However, follow-up methods are used because holistic reading on a 6-point scale with a 2-point differential gives only rough placement. Also, the test is not intended to pick up subtle gains in writing skills.

For those schools whose faculties are resistant to reading writing samples, Lyons suggested using a writing sample in conjunction with an objective test and agreeing to a cut-off point on the objective test. Some schools, like Clarkson College, are beginning to bring faculty from other departments as readers. It was also argued that for placement purposes an objective test such as ETS self-scoring test can indicate if a student has sufficient sentence-level skills to be taught writing skills.

There was some discussion about both the advantages and the dangers of calling in a professional organization such as ETS despite its expertise. When seeking outside advice, faculty and administrators should look for models and information at schools that are comparable to its own. As a beginning, Lyons suggested conferences should examine the preliminary summary of college testing programs prepared by NTSW. A final version of this data will be mailed to conference participants.

Marsha Cummins, Recorder
Bronx Community College

(Continued on page 6)
WHAT ARE WE TESTING FOR?

Moderator: Richard Donovan, Co-Director, National Testing Network in Writing

Panelists: Daniel Fader, University of Michigan
          Alison Bernstein, The Ford Foundation

Alison Bernstein introduced the issue by giving five reasons for the continued testing of writing in colleges. First, the testing of writing "legitimates" those of us who teach writing. Second, we test in order to retain our jobs. Self-interest dictates that testing is a useful way of determining and demonstrating what it is we are accomplishing. Third, we test to implement a legislative mandate that teaching effectiveness be examined in the light of student performance. Fourth, we test, as in triage, to sort out the mortally wounded from the salvageable, to discover better prepared students and those who need more help, and to discover better ways of using university resources. Finally, we test to diagnose students' needs. Testing is a tool to give us more information about students' needs so that we can fashion programs that will enable them to develop their writing skills.

Bernstein ended by cautioning us to use testing wisely, noting that tests can be dangerous weapons in the wrong hands: how we use tests and who controls them may determine the effectiveness of our writing programs.

Dan Fader, in his opening remarks, said that at the University of Michigan the testing of writing has one main purpose: to see if students are prepared to survive in classes at the University. If students cannot write, he claimed, they cannot survive in college. The writing program at Michigan recognized the admonitions from the rest of the faculty that student writers have to be able to write correct grammar, to argue coherently and logically, and to demonstrate a sense of organization and order. The writing test at Michigan measures these faculty goals in student writing, and their writing program is tailored to the same goals.

Most of the discussion that followed centered on Fader's presentation. The preoccupation with testing students' preparation for survival at the University of Michigan raised several questions from the audience: Does personal meaning, purpose, and investment in a piece of writing do anything to improve mechanics and organization, the latter being among the more important faculty criteria at Michigan for survival? How are students prepared at Michigan for survival in the "outside" world? What about thinking visually? Working collaboratively? Having students generate theses for themselves? Are we training or educating? Are we teaching in order to have students survive in all their college courses or are they more immediate and appropriate goals? Can we change our colleagues' views on how students should write? The questions generated heat as well as light, but in sum, the discussion served to raise the consciousness of all participants, especially concerning the need for more experimental programs and more research.

Uly Kapilli and Jerry Megna, Recorders: Brooklyn College

OPPORTUNITIES FOR RESEARCH IN TESTING

Facilitator: Alice Brekke, California State University

A large and diverse group assembled to discuss opportunities for research in testing. Participants ranged from experienced researchers in settings with well-established testing programs to those just beginning testing programs, who were curious about what kinds of research are appropriate and possible.

Alice Brekke began the session by describing the large-scale California State testing program. She then discussed various kinds of research presently being carried out on data from the testing program: test analysis (i.e., item analysis, test reliability); evaluation procedures (i.e., rater reliability, hidden rater evaluation rubrics); student demographics (i.e., the effects of such variables as ESL background, ethnicity, high school background, attitudes, and major); and the effects of remedial programs (i.e., retention; on career success). Several other participants then described their own experiences with testing programs and research activities associated with them.

The following concerns were among those discussed:

- What kinds of research studies can be appropriately based on placement testing?
- What opportunities do writing programs outside of English departments offer for research?
- What special constraints exist in such programs?
- How can equivalence of tasks be assured?
- To what extent can the traditional techniques of experimental research (such as tight control of variables, random assignment to treatment groups) be applied to research based on large-scale testing?
- What differences are there between political research aimed at institutional decision-making and scientific research aimed at discovering new knowledge?

Michael G. Southwell, Recorder
York College

TRAINING READERS, CHRISTOPHER, (continued)

a strong incentive. Both readers and the institutions in which they work should be cognizant of the decisive influence the outcome of assessment will have on the experience of the student and the development of the curriculum.

Peter Michalski, Recorder
New York City Technical College