RESEARCH OPPORTUNITIES
Facilitator: Lynn Quitman Troyka, Queensborough Community College

Lynn Quitman Troyka opened the session by briefly describing a "taxonomy of purposes for research in writing assessment" that she had recently developed. Troyka explained that she had created her taxonomy in an effort to set contexts for current and future research in the testing of writing. Troyka suggested that this taxonomy can help researchers discover research opportunities because it seeks to clarify what research has been done and what research still needs to be done. Troyka noted that the taxonomy categorizes research according to three major purposes. Each purpose is then delineated with key questions.

After presenting her taxonomy, Troyka asked each person at the discussion session (N=35) to introduce himself/herself, to share with the group whatever research question she/he was presently interested in. Troyka also asked each person to try out the taxonomy, in part to use the taxonomy as an organizing focus of the discussion and in part to help flesh out the taxonomy. Finally, as the discussion began Troyka urged people to take notes about other people's questions so that the network concept of NTNW could be carried on beyond the days of the conference. The taxonomy follows:

I. Research to Confirm
A. Does the test reflect the curriculum?
   1. Is it an essay or multiple choice test?
   2. What essay topics are appropriate?
B. Does the test's units of measurement truly measure statistically what teachers want to be measured (that is, does the test drive the statistics or do the statistics drive the test?)
C. Does the test adequately describe the population?

II. Research to Monitor
A. Is the test being administered correctly (setting, timing for its target population, security, etc.)
B. Is the test holding steady over time? (e.g., the CUNY audit)

III. Research to Explore
A. What results are derived from essays written in various modes of discourse?
B. What time variables are best for essay tests? (20 minutes to 2 days to x)
C. What scoring innovations are possible?
   1. What possibilities are there for scoring essay tests other than holistic, primary trait, or analytic?
   2. How can "gain score" adjustments be made for essay test scores?
D. How do student attitudes relate to test performance?

EVALUATING A TESTING PROGRAM
Facilitator: James Davis, Monroe Community College

In his opening comments, James Davis focused on the purposes of testing programs. He reminded the group that any evaluation must be based on the original objectives of the program. He then discussed the testing program evaluation studies being conducted at Monroe Community College.

On the basis of data collected from Monroe's writing competency tests and placement decisions, Davis raised the question of whether elaborate placement/testing procedures are actually necessary. What, he asked, is the justification for designing placement tests and an elaborate structure for remedial instruction if students who are placed in remedial courses do not seem to benefit from these courses? Davis described the Monroe study testing the hypothesis that students who are designated as "remedial" students by the placement test would do just as well in regular freshman composition courses as in remedial courses.

The results of the experiment suggested that taking remedial English rather than Freshman Composition was actually counterproductive to students—students in the regular course did as well or better (in terms of grades and attrition rates) than did students in the remedial course. Davis believes that these findings indicate a failure of remedial instruction even though the reasons for such failure are not easily identified.

The participants questioned the validity of Davis' conclusion. Some asked about the experimental design. Others were unsure that the data justified the conclusion regarding remedial English. Davis recognized problems inherent in any statistical analysis with so many variables, and he cautioned the participants to sharpen their critical stance in regard to evaluating testing programs.

THE POLITICS OF TESTING
Facilitator: Richard Hendrix, Fund for the Improvement of Postsecondary Education

After a brief comment on the meaning "politics" has for him ("the 'politics' I see daily are national issues...testing is not that much debated in the halls of government"), Richard Hendrix invited those assembled to introduce themselves, to say where they come from, and to state briefly what "politics" in the context of testing means to them. For a number of people, particularly those from large public instus, "politics" related to external factors, such as convincing legislators and university administrators of the value of strong writing programs. Others focused on internal issues in using the term "politics." For example, can testing provide a means of convincing faculty in other departments to support a program of writing across the curriculum? What effect does testing have on students? Does it help them master writing more quickly or only catch "losers" earlier? What effect does testing have on teaching? Does it, for instance, contradict what we have learned about the composing process? What are the societal implications of testing? Does it close off higher education to the minorities or the poor?

Much of the discussion focused on ways of gathering college-wide support for a program in writing across the curriculum. Testing can help by concretely defining levels of student ability. When testing involves a writing sample, it communicates to faculty and to students that writing is much more than grammar, and it helps clarify relationships between thinking and writing.

Another concern of the group involved the use of tests: given that tests are here, how do we make the best use of them? A number of suggestions emerged. Have faculty take the writing test to see what the experience is like and to be better prepared to deal with student reaction to (Continued on page 16)
Facilitators: Leo Ruth and Sandra Murphy  
Bay Area Writing Project

Leo Ruth began the session by reminding us that all writing tasks involve reading problems which the test maker must attend to. Sandra Murphy seconded this by further reminding the group that classroom prompts are negotiable; students may question assignments and teachers, while working to make themselves understood, disclose goals and suggest strategies. Knowledge of the teacher and the course context also helps. Writing assignment test prompts, however, must stand on their own.

With this introduction, we then got to criticize a sample real-life prompt. Some of the questions we raised involved content: Would most students know enough to write on this topic? Would they have to worry about offending the unknown reader? Most of the questions concerned the rhetorical mode and structure of the essay elicited by the prompt. The prompt was criticized for calling for three tasks: Is this too many for ease of students' performance or coherence of the product? What variations of emphasis or organization of the tasks are tolerable? If the prompt says, "Describe a problem," are we asking for a description? How will exam readers react to narrative writings which students find easier and do better, if the readers are expecting argument? These and other questions called for agreement between test makers and graders relating the nature of the prompt to the purpose of the test and an assessment of the abilities of the test-takers. As an example of the first consideration, do we want typical writing or good writing and, if the latter, what kind of good writing? As an example of the second, are our test-takers sophisticated enough to distinguish between "real" questions and "school" questions and to respond with appropriate behavior? One participant told of a returning adult student who refused to respond to her prompt on the ground that she believed the teacher knew the answer to her question.

Many of the participants, particularly those from CUNY, were interested in the discussion of "agree-disagree" prompts (CUNY uses this format for its Writing Assessment Test). Though research suggests the superiority of this type of prompt, participants pointed to problems. The phrasing of the question may give agreeers an edge, since the question offers suggestions on what to say. For the insecure writer, agreeing generates more verbiage, while disagreeing, which permits and even enforces a narrower focus, may generate better essays from those for whom verbiage generation is no problem.

Besides the sample prompt, participants were given two other handouts: "Guidelines for Developing Topics for Writing Assessments" and "An Evolving Model for Studying the Writing Assessment Episode," the latter detailing relationships among participants (test-maker, test-taker and test-rate), processes and products. The handouts turned out to have anticipated many of the participants' questions and concerns. Additional handouts can be obtained from Professor Leo Ruth, Bay Area Writing Project, University of California, Berkeley.

Charles Pitch, Recorder  
John Jay College

POLITICS OF TESTING, (continued)

the test. Have faculty generated the criteria for grading a test, then have them participate in grading. Realize that there is significant public support for testing. One participant described a junior-level reading/essay test which avoids the pass/fail grade and asks instead, "What will this student's program for the next two years look like?" Students at this institution also have a chance to revise their essay the next day, and faculty from all disciplines are involved in grading the tests and counseling students.

Are we observing a natural cycle that begins with public alarm at poor writing and that leads to a demand for testing, new courses, refinement of the test, and ultimately an impetus toward programs in writing across the curriculum? Although some participants felt that the notion of a cycle accurately describes recent developments on their campuses, a participant from a SUNY campus argued that there is too much rigidity in the system to permit the cycle to occur. Another participant pointed out that if a movement toward writing across the curriculum is to develop, it must be actively promoted from the very top, preferably by the president of the college.

Hendrix remarked that the discussion confirmed his hypothesis that testing has been a divisive issue on campuses. On the one hand, assessment has been used to get students, especially adults, into education, to develop alternative routes for education, and to educate by defining competencies. On the other hand, public outcry for a return to basics, mandated state-wide or city-wide tests, and new research interests in testing has lead to more and more tests, some of which screen students out of education. Thus, he concluded, we stand at an interesting and complex moment in the politics of testing.

Nancy Black, Recorder  
Brooklyn College

TRAINING READERS, STANLEY, (continued)

reading decides the issue. Some participants, mindful of budget constraints at their institutions, wanted to know how to keep third readings to a minimum. Others wanted to know who should shoulder this burden in addition to their other administrative tasks? Should special, experienced, "third readers" be designated (or elected)? What should chief readers do when they "know" that a third reading is incorrect? Should some of the third readings be discussed and reviewed at some point in the testing session? What procedures are used for dealing with consistently aberrant readers?

There was a general consensus that many of the problems associated with essay readings could best be resolved by developing good training sessions for readers. Such sessions should be repeated each time a reading is held, and readers should be required to attend them. The general aim of the test administrator should be to impart to the readers a feeling that they have an active input into the testing program. Testing sessions should not be too long; should take place in a pleasant, collegial environment, and should be oriented to reader participation. The ideal goal should be to make "training" a continuing and natural, organic part of every reading, for everyone involved.

It would seem from the geographic range of the session's participants that testing programs throughout the United States (and elsewhere) are likely to grow rapidly in the near future. This expansion will be especially interesting for us at CUNY as many of these programs are based on the CUNY model. This session indicated that in the future we will certainly be able to learn more from each other.

Harry Brent, Recorder  
Baruch College