EXPLORING THE TERRITORY OF MULTIPLE-CHOICE

Science consists in measuring what is measurable, and in rendering measurable what is not. —Galileo

In her "Statement on Criteria for Writing Proficiency" (1976), Mina Shaughnessy concisely maps out two "distinct territories of competence" that pertain to the assessment of writing:

One territory we can call the territory of choices, which is concerned with the quality of decisions a writer makes in the selection of words and sentences and the structural and rhetorical devices used. The other territory we can call the territory of given, which is concerned with correct forms. In the first territory a writer can be judged to be persuasive or unconvincing, interesting or dull, precise or imprecise, organized or disorganized, etc. In the second territory he is right or wrong, according to the conventions of the written code; that is, his grammar, his spelling, his punctuation, or his word choice will simply be perceived as right or wrong by the general reader.

Though Shaughnessy is interested mainly in establishing criteria for holistic scoring essays, her two territories of competence seem to me useful for thinking about writing assessment in general. Should not all tests of writing proficiency, no matter what the format, take into account the broader areas of structure and strategy as well as the territory of correct forms?

Correctness at the sentence level—the "territory of given"—has been, of course, a much more congenial terrain for yielding objective test scores than has the rugged region of choices. That is why objective tests of writing are essentially tests of grammar, capitalization, spelling, punctuation, and the like. In fact, most standardized tests systematically separate Shaughnessy's two territories; test designers customarily package all sentence-level problems in the objective part of the test and assume that the structural and rhetorical areas will be covered in the writing sample. This "solution" works out more or less well (depending on the quality of the test and the training of the readers), though if we conducted an informal survey we would probably discover that instructors largely tend to ignore the multiple-choice tests, which they view as merely an artificial assemblage of discrete subskills, while the test specialists generally belittle the reader-scored essay samples, which they regard as unreliable and expensive "window dressing" that may make the teacher feel better about testing but which adds no significant data to the assessment.

This professional disagreement is surety at the center of writing assessment problems today. It would be a step in the right direction if NTNW could encourage writing specialists to think more seriously about the statistical realities of educational measurement and test specialists to consider more closely the ways people actually write. For too long, test specialists have concentrated their attention on the smaller areas of writing rather than the larger, more fundamental areas of contemporary teachers of composition. Because discrete grammatical and mechanical problems are quite easily tested in objective formats, writing tests have been almost exclusively "handbook" tests. As a result, the professional assessment of writing has seriously lagged behind both the most fruitful and promising work done in composition over the past decade or so. Most standardized writing tests, it is sad to note, are designed as though discourse theory, generative rhetoric, and current cognitive approaches to composition have no relevance to the evaluation of writing proficiency.

Yet, writing specialists themselves are partly to blame for the inadequacies of standardized writing tests. For one thing, writing teachers in general have made little effort to understand even the rudiments of educational measurement. For another, they have so committed themselves to the reader-scored writing sample as the only legitimate method of writing assessment that they have, in effect, closed the door on new developments in objective testing. Because teachers of composition have complained so persistently about objective tests per se, there has been an apparent incentive for test specialists to improve such instruments. The habitual complaint about objective tests is, of course, that they do not examine a student's actual writing, in other words, the tests lack "validity," or what test specialists refer to as "face validity." But a good testing procedure does not necessarily have to duplicate precisely the skill being tested. Instructors should remember that tests are by their very nature synecdochic; we measure a part in order to assess the whole. To think otherwise would place us in the absurd situation of Borges' exacting cartographers who could not be intellectually satisfied with a map of the Empire until it was as large as the Empire itself.

I am not—I should add—arguing for the discontinuance of the writing sample, though that is not as crucial as it once was. The fact is that the"writing sample" is not a properly designed evaluation tool, and the cartographers' "perfect" map. I do think, however, that writing specialists should be more receptive to the possibility of designing an objective test that will better reflect their professional interests. Face validity is not everything. Objective tests, as measurement specialists well know, are often highly reliable, accurate, convenient,

(Continued on page 19)

Anne Ruggles Gere

One area which does not often receive immediate consideration from evaluators of writing is attitudes toward writing. For example, Charles R Cooper and Lee Oddell's Finding Writing: Describing, Measuring, Judging (Urbana, III, NCTE, 1977), a best seller which contains some of the best professional thinking on the subject of evaluation, does not mention evaluating attitudes toward writing. Other professional literature on evaluation emulates the patterns of Evaluating Writing: methodologies for scoring, discerning intellectual growth and identifying syntactic features appear frequently, but attitudes toward writing are largely ignored by evaluators.

The profession's lack of attention to attitudes toward writing is mirrored by lay people concerned with the quality of student writing. I have been involved in evaluating writing programs for both secondary and higher education and have participated in discussions with administrators, parents, and teachers from other disciplines. Words such as "competence," "intelligence," and "proficiency" figure prominently in such discussions, but the word "attitude" is rarely uttered.

Attitudes represent a complex set of cognitive, affective and behavioral features. That is, attitudes include what one thinks, how one feels, and what one does with regard, in this case, to writing. Accordingly, questions probing attitudes toward writing can range from "studies grammar improves writing skills" (thinking) to "I like what I write" (feeling) to "I leave notes for members of my family" (doing). The Emig (1979) scale contains items which are predominantly behavioral; Gere