

## EXPLORING THE TERRITORY OF MULTIPLE-CHOICE

*Science consists in measuring that which is measurable, and in rendering measurable that which 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 sentence patterns and rhetorical strategies. The other territory we can call the territory of *givens*, 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 choices will simply be perceived as right or wrong by the general reader.

Though Shaughnessy is interested mainly in establishing criteria for holistically-scored 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 givens"—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 surely 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 those areas of writing *least* interesting to 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 little 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 plausibility, 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 their 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 instrument can at times seem as cumbersome as the cartographers' "perfect" map. I do think, however, that writing teachers 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)



## ATWAN (Continued)

economical, and valid. They also correlate very highly with data obtained from writing samples. Much can be learned about student writing from objective testing methods, especially if problems and patterns of problems can be isolated for diagnostic use and reliable difference scores can be obtained to measure growth. What is sacrificed in face validity is frequently gained in coverage, diagnostic pinpointing of problems, information about group characteristics, and statistical reliability—important matters for English departments seeking maximum information about a large group's proficiency at a minimum of cost and time.

Given the many advantages of objective testing—assuming that reliability, coverage, convenience, and economy are no small matters—and given the current importance of rhetoric and discourse theory in composition, I would like to suggest that NTNW explore the possibility of devising an objective test to measure the area of writing that Mina Shaughnessy so aptly calls the “territory of choices.” Such a test would examine a student's proficiency in organization and development of ideas, control of tone and sense of audience, problem-solving, logic, and inference. It could also yield useful information about a student's ability to handle narrative patterns, argumentation, classification, causation, definition, and the like. I am confident that these large elements of discourse could be accurately and sensibly broken down into concrete operations and “testable” choices; surely, we can tell when a causal argument is fallacious, when an analogy falls apart, when a comparison is faulty, when a deduction is invalid, when an item in a series is misclassified. The advantages of a rhetorically-oriented composition test over the conventional handbook tests should be clear: it would eliminate the standardized labeling and editing exercises

that writing instructors consider pointless; it would ask the student writer to work within a broader range of writing choices; it would possess greater face validity and credibility; it would provide instructors and departments with convenient diagnostic information about a wide variety of writing skills; and, perhaps, most important, it would responsibly convey to the student a more realistic sense of what English composition is.

To the best of my knowledge, no objective writing test has systematically measured how well students perform in the territory of choices. A few tests (most notably the CLEP English Composition) have tried to include a broader range of problem-types, but none of them takes sufficiently into account rhetorical modes and patterns or problem-solving strategies. Parts of these published tests may, however, provide writing and measurement specialists with some prototypes to start with.<sup>1</sup> I am hoping that one of the accomplishments of NTNW will be the development of (or at least the delineation of criteria toward) an objective writing test that would satisfy both the measurement profession's reasonable demand for reliability and the writing profession's equally reasonable demand for plausibility. Between these two regions is a wide-open territory well worth staking off.

<sup>1</sup> In my *Random House Tests of English Composition* (1977, 1980), I included a few organizational, logical, and tonal problems, though the tests remain essentially “handbook” tests.

*Robert Atwan has taught at Rutgers University, served as an Associate Examiner at Educational Testing Service, and been a consultant to several national testing programs; he is author of the Random House Diagnostic Tests of English Composition and is co-editor of two composition books, Popular Writing in America and Thinking in Writing.*