
QUALITY ASSURANCE STANDARDS IN WRITING ASSESSMENT

Speakers: Sybil Carlson, Educational Testing Service
Mary Fowles, Educational Testing Service

Introducer/Recorder: Karen Greenberg, National Testing
Network in Writing

As the field of essay testing develops, policies and procedures are being established, but important issues remain unresolved. Sybil Carlson, a researcher, and Mary Fowles, a test developer, discussed some of these unresolved issues and reviewed current trends in essay testing.

Carlson began by explaining that her perspective on standards for the direct assessment of writing, developed through her research on instruction in problem solving and in the open-ended assessment of performance is more appropriately described in qualitative than quantitative terms. However, she noted that qualitative assessments must also be valid and reliable. Indeed, because approaches to essay testing can vary considerably in quality, Carlson stated that it is important to develop and apply guidelines that will assure us that the information obtained through the assessment is relevant and accurate. Carlson then pointed out that although the American Psychological Association (APA) standards serve well as guidelines for objective testing, they are not directly transferable to direct assessment, nor do they provide sufficient detail in areas that are more critical to direct assessment (such as scoring).

She then outlined some of the substantive issues that our profession still needs to address. First, we need to define more clearly and comprehensively the writing competence that we are attempting to measure. Next, we must determine the extent to which performance is equivalent, in the several instances in which it might vary (i.e., across topics, tasks, and population groups). Third, we must make our criteria explicit as the definition of competent writing is translated into assessment practices. Before assigning labels to a possible universe of writing features that contribute to the total effect of a piece, Carlson said that we need to know the extent to which these features can be identified and evaluated independently. She concluded by noting that as we

reach some agreement about a set of common guidelines or standards for direct assessment, we will be able to communicate in a language in which we share common definitions.

Mary Fowles then presented the ETS Guidelines for Developing and Scoring Free-Response Tests. She discussed the steps involved in planning direct writing assessments and in developing writing test specifications. Next, she explained the guidelines for writing the scoring specifications and techniques for developing writing test items and scoring criteria. She also described methods for pretesting the test items and field criteria, and cautioned conferees to ask questions about the test and the scoring criteria, such as "How well do the examinees understand what they have to do?" "Do the tasks elicit the responses that were expected and desired?" "Can the criteria and scale be used to score the test easily?" "Are the readers using the scoring system in the way that it was intended?" "To what extent do the readers agree on the scores that they assign to the responses?"

Fowles then described each of the steps in administering and scoring an essay test of writing. Next she explained the use of statistics to evaluate the test and the scoring system. She ended with a discussion of techniques for evaluating the validity of writing tests.◇