

SELECTING THE BEST ESSAY TOPIC

Speakers: Gertrude Conlan, Educational Testing Service
Pamela Chambers, Educational Testing Service

Introducer/Recorder: Albert Porter, Mercer County
Community College, N.J.

Gertrude Conlan and Pamela Chambers described the procedures used by Educational Testing Service (ETS) to select and field test essay topics for two standardized testing programs: The New Jersey College Basic Skills Placement Test (NJCBSPT) and the College Board English Composition Test (ECT). The NJCBSPT is administered to 50,000 New Jersey College freshmen each year to help identify those needing remediation. In December of each year, the ECT is completed by 90,000 college-bound high school students. Students' essays in both programs are scored holistically at ETS by trained faculty readers under controlled conditions.

Although the audiences for the two programs are quite different, the procedures used to select topics and the characteristics of the most successful questions are remarkably similar. Committees of English faculty generate an initial pool of 50-100 topics from which ten or more are selected for field testing. Using guidelines from ETS, each topic is pretested in the classroom to obtain a sufficient number of writing samples from students representative of the target population. Readers consider the success of each topic in prompting students to develop their ideas, form arguments, and defend their opinions. The best topics are accessible to students, allowing them to show how well they can write.

Typically, only a few topics survive the pretest reading. If a topic produces inflammatory responses or fails to differentiate between weak and strong writers, it is discarded. Some topics are judged inadequate because they generate uninteresting writing or reveal unintended interpretations of the question. Even the best topics from the pretest stage may fail to work after adoption. Unexpected events immediately preceding the test may disproportionately influence students responses to a question. For example, a major world event or a widely viewed television show can homogenize students' viewpoints and arguments on a topic. Obviously, the field testing effort cannot control for such occurrences.

Using illustrations from actual topics and student responses, Conlan and Chambers demonstrated the features of topics that have succeeded and failed. Some topics judged marginal during pretesting succeeded after recasting to clarify their meaning or amplify their intent. For example, the replacement of vocabulary found to be obscure to certain cultures (or reflecting regional differences) has made some

marginal topics more accessible to a broader range of students.

In their closing remarks, the presenters emphasized the most important considerations in topic selection. To the extent possible, topics should be:

- clearly stated, using language familiar to the audience
- appropriate to the population being tested
- reasonable for the allotted writing time
- suited to the purpose of the test
- accessible to the broadest range of students
- interesting to write about
- capable of producing writing across the range of student abilities

Some participants were concerned about the limitation of twenty minutes for writing the essay, particularly in light of instructional emphasis on the importance of the prewriting, drafting, composing, and revising processes. Although the presenters acknowledged the potential difficulty of the time constraint, they noted that it would be impractical in a testing situation to provide students with the necessary time and distance to simulate actual writing exercises. In a separate project now nearing completion, ETS is analyzing the influence of additional time on the quality of writing produced by examinees taking the NJCBSPT essay. Under controlled conditions, students in that study were given twenty minutes, thirty minutes, and twenty minutes with ten minutes of prewriting; results of the investigation will be available in late spring.

A separate question involved the comparability of scores obtained using different topics. The presenters warned against such direct comparisons, even when scores are derived using trained readers and carefully selected rangefinders. However, ETS is studying both the inter-reader reliability across different scoring sessions involving the same topic and the comparability of scores when the same topic is repeated in separate test administrations. These studies primarily focus on the effectiveness of reader training.◊