
CONDUCTING RESEARCH FOR A NATIONAL ESSAY TEST

Speakers: Richard Swartz, GED Testing Service
Wayne Patience, GED Testing Service

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

Richard Swartz began the session by asking the following question: How do you design an essay test that can be administered virtually any day of the year and scored-- according to the same standards-- in one of several hundred locations throughout the United States and Canada? This is the question that the researchers who developed the new essay section of the GED Writing Skills Test had to answer.

Swartz noted that the current Writing Skills section of the GED includes multiple-choice questions intended to measure a person's ability to use Standard Written English clearly and effectively. The essay section of the GED test will be

administered beginning in 1988. The GED created a Writing Committee whose responsibilities included reviewing and evaluating potential essay topics and selecting rangefinders and training papers. The Committee gave considerable attention to selecting the best possible method for scoring the writing samples obtained from high school students, and holistic scoring was selected as the optimal method. Swartz pointed out that holistic scoring is a quick and relatively simple procedure, so it is an efficient method for scoring large numbers of essays. In addition, holistic scoring offers the most accessible method of attaining high degrees of inter-rater reliability, and these were necessary in order to answer the primary research question involving the correlations between essay and multiple-choice scores. In order to control the drifting of standards from one reading to another, a scoring guide was developed for each topic. Because the guides were created inductively--by describing the common characteristics of agreed-upon model papers--they could be regarded as descriptive rather than prescriptive.

Wayne Patience then discussed the evaluation of potential essay topics. He noted that in order to minimize the problem of controlling topic variability, it was decided to restrict GED essay topics to one type--expository. A sample topic is as follows (the time allotted is forty-five

minutes):

The automobile has certainly been responsible for many changes in the United States. Some of these changes have improved our lives and some have made life more difficult or unpleasant. Write a composition of about 200 words describing the effect of the automobile on modern life. You may describe the positive effects, the negative effects, or both. Be specific, and use examples to support your view.

After potential topics are reviewed and edited by the GED Writing Committee, they are field-tested in administrations to both high school seniors and GED examinees. If topics survive these evaluations, statistical evaluations are performed (i.e., a Kruskal-Wallis test of the homogeneity of score distributions and a test of the equality of correlations between multiple-choice and essay scores).

Both Swartz and Patience noted that one of the most important conclusions gained from their experience is that constant vigilance is required to maximize essay score scale stability and reading reliability. They concluded their session by stating that the development of the GED national essay testing program posed unique problems in large-scale writing assessment, and some of the solutions may apply throughout the field of writing assessment.0