

ESSAY TEST TOPIC DEVELOPMENT

Speaker: Gordon Brossell, Florida State University

Introducer/Recorder: Bruce Holland, University of Akron, Ohio

Gordon Brossell posed the question, "What do we know about essay topic development for large scale assessment?" at the outset of his session and proposed to answer it in light of recent research and practice. He first summarized the following maxim, principles, and guidelines relevant to making assignments for writing assessment, set forth in an ERIC report written by Leo Ruth and others in 1982.

Maxim: A well-framed assignment for writing assessment provides an occasion for writing in which students' powers of expression and communication are stimulated to their maximum abilities.

Principles:

- 1) Subjects chosen should potentially be interesting to the teacher/reader.
- 2) Subjects chosen should potentially be interesting to the student/writer.
- 3) Data furnished should be open to presentation in several forms: e.g., statement, question, picture, cartoon.
- 4) The assignment should be meaningful within the student's experience.
- 5) The assignment should elicit a specific response and limit content or form or both.

Guidelines: Questions to consider in creating topics for writing assessment are:

- 1) What is the assignment's purpose (e.g., to ascertain placement, minimum competence, etc.)?
- 2) What will be done with the results?
- 3) What is the nature of the student population being assessed?
- 4) What is the most appropriate topic/subject for this group?

Brossell next identified three major situational variables in large-scale writing assessment:

- 1. Topic Variables:** vocabulary and syntax; subject matter; the assumed or intended mode; rhetorical specification (i.e., information about purpose, audience, situation, etc.).
- 2. Writing Variables:** topic interpretation; writer apprehension and anxiety.
- 3. Procedural Variables:** instructions on the exam; time limit; scoring systems and readers; test environment.

Brossell discussed the effects of changes in vocabulary and syntax on writing assessment results. He noted that although common sense suggests such changes will make a difference to the writer, recent research has called this view into question. In Karen Greenberg's 1981 study at CUNY, variations in the cognitive demands and the experiential demands of a writing assignment were found to produce no significant difference in the essay scores. Gordon Brossell's own study, done with Barbara Ash, showed that neither wording changes nor subject changes had much effect on an essay's score. Brossell cautioned that he was not claiming that differences in subject are not important, but that we can't predict with certainty the degree and kind of change that will produce

a different effect. The problem of subject-matter variation has been investigated by Thomas Hilgers of the University of Hawaii. According to Hilgers, the fact that some content in a writing prompt is more familiar to some students can be dealt with by providing more information beforehand to overcome any deficiencies individual students might have. Brossell claimed, however, that carefully constructed prompts could preclude this problem. The mode of discourse variation involves such questions as whether argument is more difficult than narrative: Aviva Freedman and Ian Pringle's 1981 study found that twelve and thirteen-year-old students were able to recognize the conventional schema for a story more easily than for an argument. And despite the belief among sentence-combiners that writing is judged better if it is syntactically more complex, the research results are mixed: the syntactic complexity position has been called into question by James Moffett and Michael Holzman, and several studies show vocabulary to be more significant than syntactic complexity.

Rhetorical specifications are another source of topic variation. The amount of information given to a writer may help him or her focus and organize the essay on a timed test. Brossell found, in a 1983 study, that a moderate amount of information seemed to help writers more than either very short or very long, fully specified prompts. Another study, done by Stephen Witte and others, found that the highest scored papers came from prompts that had high audience and content specification but low purpose specification. These findings, said Brossell, cut against the expectations of writing teachers, but may indicate only that the teaching and the testing of writing are much more divergent enterprises than they appear. One of the most comprehensive research projects relevant to the issue of rhetorical specification is Alan Purves's study for the International Association for the Evaluation of Educational Achievement. Purves has created a classification system setting forth fifteen dimensions of writing assignments:

- | | |
|-----------------------------|-------------------------|
| 1. Instruction | 9. Tone, style |
| 2. Stimulus | 10. Advance preparation |
| 3. Cognitive remark | 11. Length |
| 4. Purpose | 12. Format |
| 5. Role | 13. Time |
| 6. Audience | 14. Number of drafts |
| 7. Content | 15. Criteria |
| 8. Rhetorical specification | |

These fifteen considerations can provide teachers, test makers, and researchers with a set of tools for adjusting writing assignments. Brossell noted Purves's own caution that the schema is highly tentative, but also added that it provides an excellent starting point for identifying and characterizing assignment variables.

On the subject of writer variables, Brossell limited his discussion to topic interpretation. Each writer's understanding of a writing assignment, said Brossell, may be different. Leo Ruth and Sandra Murphy offer four generalizations about the interpretation of writing tasks:

1. Different writers have different notions about the significance of certain features of a topic.
2. Writers and evaluators don't always interpret a topic in the same way.

3. Writers may construct different writing tasks for themselves at different stages in their development.
4. Differences in responses to writing prompts require further investigation.

The final part of Brossell's presentation described a procedure that he and James Hoetker have developed for generating sets of exam topics that are content-fair, i.e., topics that a great majority of students can write about. As part of a project to supply sets of topics for Florida's College Level Academic Skills Test (taken by college sophomores), Brossell and Hoetker first identified several criteria for composing prompts less susceptible to misreading:

1. Topics should be brief.
2. Topics should be cast in similar syntactic patterns and share semantic elements.
3. Instructions should be brief and simple and identify the real audience (essay raters).

Eventually they adopted as a "master topic" the two years of college experience just completed by the test takers. Adapting the concept of a formal definition (a class term plus differentiating criteria), they found they could construct a pattern or "frame sentence" capable of generating a set of similar topics. For example:

Class Term	Differentiating Criteria
a. A novel	that many students read/which may affect them significantly.
b. A book	that many students read/which may affect them harmfully.
c. A course	that many students take/which may affect them in important ways.
d. A habit or belief	that many students acquire in college/which affects their life in important ways.

By substituting a different noun phrase or a different set of distinguishing criteria, another set of related topics could be generated.

Brossell and Hoetker field-tested their topics twice. One optimistic interpretation of the results in the first field test was that the topics helped weaker students perform better than they might have with other topics. Student responses about the test topics were also quite favorable. Brossell found that the openings of many essays written for the field test began more purposefully. He quoted a number of summary reports from the test administrators in the second field test testifying to the efficacy of the topics used. The testimony cited the fairness of the topics for highly diverse examinees, the freedom students had to organize their responses as they wished, the interest of the topics, and the variety of responses made possible by the topics, which helped prevent boredom in the readers. Overall, there were no noticeable negative effects, though Brossell pointed out several examples of topics that "misfired." The topic-generating procedures seemed to produce an infinitely large pool of comparable topics, which helped students to identify a subject they were equipped to write about and assisted them in finding a focus for their essays. A copy of Brossell's study, "A Procedure for Writing Content-Fair Essay Examination Topics for Large-Scale Writing Assessments," is available from the author. ■