NEW QUESTIONS FOR EVALUATORS OF WRITING

Over the past twenty years, we have learned a great deal about ways to evaluate student writing reliably, fairly, and efficiently. However, there are three areas to which we still pay far too little attention.

The first of these areas is the selection of topics for students to write about. Most evaluators take some pains to identify topics that students will find engaging and challenging. But we give too little thought to this sort of question: What are the conceptual demands of the topics we assign? What kind of thinking do the topics require? What strategies must writers employ to formulate their ideas, feelings, or reactions concerning a given topic? The importance of asking these questions becomes especially clear when we consider a writing assessment that allows students to write in response to either of these assertions:

"Much of the instruction that goes on in public school classrooms does not adequately prepare students for college."

"There should be a mandatory jail sentence of at least one year for any person convicted of possessing more than three ounces of marijuana."

These assertions (and the topics they raise) seem similar in at least two respects. Both statements invite the request to "agree or disagree with this statement." Moreover, these statements might provoke a strong response from a writer. But the two topics differ in at least one important way. The first raises questions of fact. To express agreement or disagreement, a student would need to recall specific school experiences and determine whether the information/knowledge resulting from those experiences is consistent with information/knowledge required in college classes. By contrast, the word should in the second assertion raises not merely questions of fact but questions of principle, morality, and legal precedent.

Classical rhetoric makes the distinction between questions of fact, definition, and degree. And James Moffett has made a strong argument that writing about what might or should happen is quite a different matter from writing about what has already happened. All of us know of these distinctions but too frequently we fail to consider them in assigning topics. Consequently we undermine at least two of our purposes in evaluating students' writing. If students are doing substantially different writing tasks, it seems unfair to rank order their writing; we cannot say that one student's writing is more or less skillful than another's unless students are writing about topics that require similar skills. Further, we cannot determine whether student writing is improving if the demands of the topic we assign early in the term are significantly different from the demands of a topic we assign late in the term.

We cannot guarantee that all students will perceive a given topic exactly as we do. But perhaps we can avoid making students' lives needlessly difficult if we try to determine whether different topics make comparable demands of students. To make such a determination, we might follow several procedures. If we have made a practice of allowing students to choose from among several topics, we might review papers from previous years, asking: Do students consistently choose some topics and ignore others? Do students who choose topic X consistently receive higher grades than do students who choose topic Y? We might also compare and contrast some of the most successful papers about two or more different topics. Better yet, we ourselves might try to write on each of these topics. But whether we examine students' writing or our own, we need to ask such questions as these: Do different topics require writers to draw upon different sources of information? Does one topic require writers to be particularly conscious of their own or someone else's assumptions? Does one topic invite chronological thinking whereas another invites analogical thought?

A second issue to which evaluators are likely to pay too little attention concerns audience and purpose. With some exceptions, evaluators simply identify topics for students to write about and ignore such questions as these: What rhetorical purpose are students to achieve in their writing? What are the characteristics of the audience they are to address? One response to the first question is to claim that we have made purpose reasonably clear when we ask students to "agree or disagree" or to "explain your point of view and support it with evidence." But sometimes when we ask students to "agree or disagree" with topics about which they feel very strongly, they write expressive discourse rather than the persuasive discourse which we may have hoped for. That is, they may articulate their own views (perhaps using emotional language and presenting only those facts that support their point of view) while failing to enlist the audience's sympathy, establish some common ground with the reader, or anticipate the audience's objections to what the writer is saying. My point, then, is that we need either to give students explicit information about their purposes or help them specify clearly their purpose in writing. Both the writer and the evaluator need to know whether the writer is trying to convey information, to express his/her reactions or personal conclusions, or to change the reader's thinking about a specific topic.

In addition to a clear sense of purpose, both the writer and the evaluator must understand the characteristics of the audience the writer is addressing. As was the case with purpose, we have our reasons for not indicating the characteristics of the audience students are addressing. Students, we tell ourselves are writing academic discourse and they should know that any reader of such discourse has certain expectations. Or more specifically, we may argue that students know they are writing for English teachers and that they should accommodate the expectations of that audience. In some respects, these claims are entirely reasonable. An academic audience will (or should) have certain expectations about the conventions of standard written English. And most academic audiences would accept the general principle that student writing should reflect clear thinking. But different academic audiences may vary widely in the criteria by which they judge students' writing.

About some subjects, a reader may know a great deal and may, therefore, not need or may even be irritated by certain kinds of elaboration. For other subjects, about which the reader knows very little, the reader may need a great deal of background information and elaboration. A similar variation may also appear in a reader's biases on a given topic or the reader's attitude toward the writer. Consequently, both writer and evaluator must have some shared answers to these questions: How much does the intended audience know about the subject at hand? What sort of biases might the reader have? What is the reader's attitude toward the writer?

One final matter that evaluators tend to overlook is this: How do procedures for evaluating writing relate to the ways writing is actually taught? Unless writing
COMPETENCY TESTING AND WRITING PROGRAM DEVELOPMENT

Four years ago Anderson College began testing freshmen to place them into appropriate writing courses and to assess their competency when writing courses were completed. One-half of our faculty members, representing seventeen academic departments, have served as readers for the entrance and exit essays our students write. Although we initiated testing to evaluate students, we found that test results demanded a re-evaluation of course designs and teaching strategies. In fact, group testing can have serendipitous effects on curricular reform and faculty development.

Because we hoped to establish a standard of writing competency that would reflect the consensus of our faculty, we adopted a test designed by Rosemary Hake and David Andrich. 1 This method is uniquely suited for use by a diverse group of readers. Preparing for the test requires all readers to evaluate a common set of essays selected from a group; responses to these writers are the basis for computer-assisted analysis of reader differences. First, readers identify flaws in organization, syntax, usage, and mechanics. Next, they make an independent holistic judgment about the essay's overall competence. The test design predicts that even when using common criteria, readers will differ in their assessment of a writing sample. Computer programs provide data to compensate for such differences.

An analysis of the differences among readers prompted changes in curricular design. First, we found that English faculty do not agree on tasks that differ from those of faculty in other disciplines. However, the extremes of "harsh" and "lenient" evaluation occurred among writing teachers: students were more likely to encounter divergent standards within the writing program than across the curriculum. Therefore, writing faculty now work together to articulate course objectives and evaluation criteria. Further, a student's writing is evaluated by a reader other than his or her teacher twice during each term. These measures have tempered extremes and promoted greater consistency in the program.

Second, realizing that faculty from many disciplines expressed similar values about writing relieved stress within the writing program. Recognizing our colleagues' demonstrated ability to contribute to students' development as writers, we no longer view freshmen writing courses as the last bastion in a war against illiteracy. For example, students of beginning French can earn an "A" without mastering Voltaire, as students of golf can earn an "A" without breaking ninety, but students of writing often find that the "A" is earned for nothing short of a feature for The New Yorker. As a result of a group testing, our curriculum now reflects stages of progress built upon achievable goals.

Finally, group testing helped to determine what those goals should be and how they should be ordered. We learned that as readers we define competency in terms of careful organization and logical development of ideas, not in terms of absence of error. Our introductory writing course now reflects this insight. No longer a grammar drill and usage review, the course teaches students how to impose order on their ideas and how to support with specific examples. When students reach this goal, they turn their attention to surface features as the next stage of their development.

Even more significant than curricular reform has been growth among faculty as teachers of writing. First, the testing system required us to describe the features of an essay as well as to judge its overall effectiveness. But although we could approach consensus in holistic judgments, we could not agree on the characteristics of particular essays. Now our discussions focus on describing: What is required of an introduction? How much detail is enough? What does "well-organized" mean? As a result, we learn to provide useful formative evaluation in the classroom.

An important finding was that readers diligently sought out errors in poorly-developed essays but glossed over errors in well-developed essays. We had to confront what this practice implied about our teaching. Whereas the weaker writer is overwhelmed by miscellaneous marking of mistakes, the better writer cannot achieve mastery if teachers settle for minimal competency. We are learning to vary our formative evaluations to meet students' needs at different stages in their development as writers.

The objective data from each testing situation, often contrasting with our testimony, provide a solid basis for discussing what we mean by competency, how best to help our students achieve it, and to what extent it depends on audience, context, and purpose. Such discussions highlight features of good writing that do not lend themselves to measurement—forcing us not to take our "test" too seriously.

As a result of competency testing, our faculty share a common knowledge of our students' writing skills, a common understanding of the problems students face in writing for diverse audiences, and a common vocabulary in which to discuss writing. However, we have not reached consensus about competency. The project is time-consuming; our inconsistencies are embarrassing. Participation as readers demands that we accommodate new information and other points of view. But although competency testing may not be the best way to evaluate our students, it can promote a rational curriculum and insightful teaching.

1 Rosemary Hake and Jesse D. Green, "A Test To Teach To Composition at Chicago State University, ADE Bulletin (February, 1977), 31-37.

Barbara Weaver is Director of the Writing Program at Anderson College in Anderson, Indiana; she will be serving as Assistant Director of the Writing Program at Ball State University during a sabbatical leave granted for 1982-83.

ODELL (Continued)

teachers emphasize the need to adapt writing to suit one's audience and purpose, there is little reason to believe students will benefit from an evaluator's attempt to clarify audience and purpose. If students have received instruction in only, say, writing persuasively to an unsympathetic audience, there is little reason to expect them to do well on a writing task that asks for expressive writing intended for a familiar, sympathetic audience. Unless we are careful to relate the evaluation of writing to the teaching of writing, our evaluation may tell us only that students are doing poorly with tasks on which we have no reason to expect them to do well.


Lee Odell is Professor of English at Rensselaer Polytechnic Institute.