Assessment in Context:
Toward a National Writing Project Model

I am a newcomer to the world of large-scale assessment in writing and over the past year I have been amazed by the genuine concern testing people show for both students and teachers affected by their tests and reports. At the same time, I wince when I hear of the individual suffering caused by poor tests or testing conditions.

To avoid the trap of talking about large-scale assessment without talking about the context in which such tests are developed and administered, I begin this article with a brief story. The characters in this story are fictive. The "facts" of the story I have compiled from various contexts to show a progression from beginning to end.

Daryl is fourteen, a high-school freshman in a suburban school district. It is May, and he is taking an essay exam as part of a large-scale writing assessment. If he passes, he stands a good chance of graduating in three years. If he doesn't, he will be required to attend an after-school, weekend, or summer remediation course.

Daryl is in an advanced English class—not because he is verbally or linguistically at the top of his class, but because his parents thought it would be good for him and it would make it easier to get into college. Before entering this class, Daryl enjoyed writing both in and out of school, although he did not write much. He showed a natural talent for writing, but as an easygoing child he had neither the need nor the inclination to push himself to write.

At the beginning of the year, Daryl found himself in a "personality conflict" with Mrs. Dearborn, his English teacher. When he asked his parents to switch him to another class, they refused. His father believed it was time for Daryl to learn how to survive unpleasant situations: "I've had many a boss who hasn't liked me, and I've learned how to get along," he told his son. Daryl's mother did not like the idea of her son being in anything but the "best" classes. Daryl was stuck, and over the year he learned to hate his English teacher, his "incarceration in English 9-A," and writing. On the morning of the exam, he wanted one thing: to be the first one finished and out the door.

The exam Daryl was sitting for was the brainchild of the State Board of Education, and specifically Jessie Blaycock—a member of the board representing a rural part of the state. While attending a fundraiser for State Senator Trumbell, a supporter asked the senator what he intended to do about kids who were graduating from high school without being able to read or write. The senator, already on record as opposing social promotion, turned to Jessie and said that between the two of them they were going to put an end to this foolishness in the schools. Both Jessie and the senator were good to their word and managed to pass, with only minor resistance, a motion to test all of the state's high school freshmen in reading, writing, and math. Those who failed would be "provided" remediation and would have to retake the test in the next grade until they passed. If, by twelfth grade, they failed any one of the three parts to the exam, they would be denied graduation.

Producing a Test

When the State Board of Education was about to pass the assessment program, they turned to the state superintendent to make sure it could be implemented. "With adequate funding," he said, and all in the boardroom laughed. The superintendent had been under considerable pressure from the governor's office to "produce" and to hold teachers accountable. The superintendent was not about to admit that a full-scale assessment could not be carried out with existing resources.

The testing division was informed of the board's decision and searched for the least expensive test they could buy "off the shelf." They discovered, however, that available tests were no longer conceptually valid. They had been designed years ago before the recent paradigm shift in writing instruction and did not include such things as revision, group work, collaborative writing, conferencing, or writing for a variety of audiences and for a variety of purposes. Despite their protests, they were directed to select a test and pilot it. The one they chose would be graded by a corporation from another part of the country.

continued on next page
Assessment in Context

continued from previous page

News of the impending test spread belatedly to the state English teachers association which pressured the State Department to make sure the writing test tested writing. Multiple choice tests, they said, would send the message to teachers that writing did not count—only the ability to identify and correct spelling, grammar, and usage errors. The State Department assured them that writing would be on the test but that because the governor was interested in improving the state's educational standing in the country, the test would also include a multiple choice usage test with national norms. The combined time for both portions of the test would be 40 minutes—20 minutes for each part.

The effect on classroom instruction was predictable. Teachers who had been teaching grammar and usage added a two-week unit on the five-paragraph theme—a formula that would adequately prepare their students for the writing portion of the test. Teachers who were teaching writing, incorporating grammar into the writing, added grammar and usage lessons. All but a few teachers taught with one eye on the test. The few who did not believed so strongly in their teaching that they were confident their students would do well. It was no coincidence that such teachers invariably came from more affluent and better educated parts of the state.

Daryl failed the test. Until the afternoon he received the notice in the mail, he had believed himself to be an average writer—and his class grades supported this. And at first he ignored the test results. But when his parents and teachers began to treat him as handicapped, he began to believe it himself. In his mind he no longer believed he had failed a test; he saw himself as a poor writer. In fact, when he thought of his writing he no longer pictured a sheet of lined paper with his handwriting on it, but instead saw the form letter announcing his failing grade. By the time he enrolled in summer school remedial writing, he was diagnosed as having a severe writing block.

Daryl's mother was shocked; Daryl's father outraged. When they went in to see his ninth grade English teacher, she said she agreed—Daryl was not that bad a writer, but there was nothing she could do. The principal said the same, as did the local school board. Daryl's parents had always approached public education with a sense of detached support--"leave education in the hands of professionals." It never occurred to them that the test their son failed had not been initiated or approved by teachers.

In the fall, Daryl found himself in a low ability English class; all the students in that class had failed the writing test. The school had decided to use the test results to rank students for placement.

Worst Case Scenario

The above story is a worst case scenario. It is worse in degree, but not in kind. Each character is acting for reasons that are sound within his own world, but which make little sense in the world of large-scale assessment. Here, compartmentalized thinking wreaks havoc.

Without stating the obvious, allow me to indicate some of the items that make this such a sad story for Daryl and the profession.

I would guess that there are as many motives for instituting large-scale assessment in writing as there are people involved in the decision to do it. Some goals, however, are best handled in other ways. For instance, if there is evidence of rampant social promotion, and in Daryl's story there is no evidence—only the fear of the political ramifications of appearing soft on social promotion—it is best to target teachers and schools, i.e. those responsible for it, not students.

Testing, as well as promotion, is best left in the hands of teachers. To remove teachers from the decisions involving testing, including the design and rating of tests, is to weaken the very segment of education that must remain strong. In Daryl's story, teachers were not involved. When the English teachers association did become involved, it was in a position of protesting an accomplished fact. Furthermore, the very people who knew the test was not conceptually valid, the State Department testing people, were powerless to put a stop to the assessment once it was set in motion.

Students were given the test without understanding its ramifications and importance. Instruction was shifted in a last-minute, band-aid fashion to prepare for it. Students had no opportunity to take sample tests or receive outside help if instruction in their classes had been inadequate. In fact, if a group of parents had decided to sue the State Department, I think it might have found grounds for a case.

Parents, too, were inadequately informed before the test, and not provided an avenue to discuss the results with someone who had the power to determine a retest. Based on his previous performance, Daryl deserved a retest.

"Testing...is best left in the hands of teachers. To remove teachers from the decisions involving testing...is to weaken the very segment of education that must remain strong."
Portfolio Assessment

His record of success as a writer in his classes was a much more valid measure of his ability than the 40 minute test. Portfolio assessment, which measures writing over time and includes such things as writings for a variety of purposes and audiences, and in a variety of forms or modes, is a more accurate and valid measure, i.e., it measures more of what is being taught. Portfolio assessment also supports instruction and currently is consistent with the best practices of teachers.

The remediation program Daryl attended was hastily developed to meet the quick implementation of the test. Without background information or portfolios, teachers had little to go on in developing remedial instruction.

Daryl's mistake in considering the test results for his ability to write is called reification. When test results are expressed in signs and symbols, it is very difficult for the student or the school system to control reification. A more elaborate and detailed description of the writing, of course, would decrease the likelihood of reification.

I end the story with the usual fallout of reification—ranking. Ranking is questionable under the best of circumstances. In Daryl's case, it is unjust.

A NWP Model

Missing from this story is the National Writing Project and a model for the collaborative development of large-scale assessment.

At the invitation of NWP, Mary Fowles and Roberta Camp of the Educational Testing Service met with the Advisory Board this April in Atlanta. We agreed to pursue jointly the development of such a model that could be used by states or local school districts. Although the details of this model are yet to be worked out, I expect them to include at least some of the following:

1. Involvement of teachers in the decisions to test or not to test, as well as involvement of teachers in the design, implementation, rating, report writing, and remediation.

2. Workshops for parents, students, and teachers to make instruction and assessment as consistent as possible.

3. An outside evaluator or evaluation organization to at least guide in the procedures.

4. A written study of the context in which the test is given, including social, economic, educational, and political factors.

The National Writing Project cannot hope to solve or prevent overnight the problems associated with large-scale assessment in writing. But it can become a powerful voice for writing teachers who already know who is writing well, who isn't, and why. If a school district, state, or nation wants to know about the ability of students to write, it should start with those who know, and not proceed without their concurrence.

Donald Gallehr is Co-Director of the National Writing Project and Director of the Northern Virginia Writing Project. He is Professor of English at George Mason University, Fairfax, Virginia.