Alternatives to Standardized Testing

External testing is no more necessary for learning in school than for learning out of school. It does not benefit those in the classroom, who can better assess in other ways. Standardized tests exist for people outside the classroom—for administrators and the public. All they do is compare one student or school or school system with another. This serves only to create mischief. Parents have a right to know how much schools are helping their children to learn, but they can ascertain this better from seeing their children's work and from talking with them. In good learning environments, students create a lot, which means that there is a lot for parents and administrators, as well as students and teachers, to see.

The way the public actually judges schools is by real performance in the world out of school, not by test scores, which mean little beyond academic walls. The most telling fact is that even students who once scored high and got high grades in, say, math or science remember too little to apply them later when they really need them. The complaints of employers and of graduates themselves tell us more than tests do. How well can the citizenry deal with ideas, communicate and collaborate with colleagues, make sense and use of texts, vote knowledgeably, and conceive solutions to problems?

Assessment experts, it's true, are working to make external evaluation more sophisticated than the crude multiple-choice, computer-scorable tests that have always shrunk the curriculum to fit themselves. Such experts repudiate standardized tests as we have known them and claim to be able to design testing activities that will do justice to any learning goal. But tests simple and cheap enough to permit comparison and to administer universally can never do justice to the depth and complexity of what educators are calling "higher literacy" and "critical and creative thinking" or "higher-order thinking." Furthermore, the more nearly such testing activities might succeed in assessing these desirable mental activities, the more nearly they would approximate the actual real-life performances themselves, in which case there's no need for special testing circumstances, since these performances can be observed where they authentically occur in and out of school.

In other words, if students are learning by doing, by practicing the target activities themselves, then anyone can evaluate by observing daily learning, because the learning and evaluating activities are one and the same. We can assess these activities by whatever means and standards we are all judging schools in society at large when we defend or indict them. This is ideal—if, again, the learning activities

are the target activities themselves, not exercises alleged to lead to these goals. The most efficient education would never require of learners that they do anything especially for evaluation that they would not be doing anyway in order to learn. Furthermore, if special testing activity is required, it betrays the learning goals to the extent that it differs from them. And in order to accommodate the special conditions and costs of mass measurement, it must differ a great deal.

National assessment exists to embarrass schools into improvement by comparing scores. This assumes that dereliction is the problem and competition the answer. It's a crude, moralistic, negative approach. What evidence exists that the threat of getting beaten will spur and cure schools? The fact is that, as much as anything, this very authoritarian approach has demoralized teachers and principals, who simply never have had decision-making power commensurate with responsibility for the results, because tests and texts—the major determinants of curriculum—are usually selected over their heads, if not behind their backs. Perhaps the first reason schools have found improvement so difficult is that state and district legal requirements have built standardized tests into the curriculum and into textbook adoptions so that everybody has to teach to the tests and ignore both their personal expertise and the urging of their professional organizations.

The movement toward site management aims precisely to offset such top-down governance in the local districts by delegating decision-making powers to schools and neighborhoods. But the states and the federal government are neutralizing this movement by pushing national testing farther than it has ever gone before. The presidency and the governorships seem far less willing than the local districts to give power to the grassroots, perhaps because they don't have to live with, or can't see as well, the negative consequences of their efforts to control education through testing. Proponents of national assessment reiterate that participation in this competitive testing is voluntary, but they know perfectly well that when state and federal government throws its weight behind something, parents and communities will clamor to have it in ***their*** district. Actually, district and state school systems, like the individual students in them, differ far too much in far too many ways for scores to show who is and isn't doing a good job. The reasons for poor performance go far beyond mere reprehensible character. The whole idea of improving an institution by showing it up is negative and unfair. The legislators and other politicians who are adopting this get-tough policy to weed out incompetents are the very ones who have been, in effect, blocking educational improvement for decades already by decreeing assess-

ment and procurement policies that conflict with what the best teachers are trying to do.

The official argument goes that if government permits site management and parental choice, then the educational results have to be measured against state and federal standards, to protect students from local ignorance or incompetence. But standards don't have to be set by *tests* and in fact *cannot* be set by tests, because standards are ideas of excellence that will always exceed what standardized instruments can afford to measure. Whether norm-referenced or criterion-referenced, furthermore, tests must allow most students to pass and therefore must anchor learning to *low* standards—an unnecessary self-contradiction caused by the insistence on competition. When an individual's progress is measured only against his or her past, standards do not have to be pegged low enough to accommodate masses. Comparing individuals against each other hampers everyone's progress by creating distracting self-concepts. Incessantly testing students, finally, amounts to putting them on probation throughout their youth. This creates chronic problems of low self-esteem and resentment toward schools, which should be there, after all, to serve them, not to shame and intimidate them.

How can we set standards without tests? Well, where do the criteria for tests come from if not from prior notions of what to look for in growth? It is these notions of growth in certain areas that provide standards. This is why I have emphasized in this book the *detecting* of growth in language—perceptive observation. It is an effort to describe signs of verbal growth that educators and parents may look for. Most of these kinds of growth, which I have tried to summarize periodically in italicized statements, could not be transformed into acceptable tests even were that a good idea.

Furthermore, some kinds of growth occur in the parts of discourse—like vocabulary or sentence structure—so that if tested in isolation, as is the traditional practice, would perpetuate exercises with them in isolation. Teaching to tests inevitably causes learning activities and conditions to resemble testing activities and conditions. But an observer can *notice* how a student is developing in the substructures of discourse without isolating these in the student's mind. In fact, such *in situ* observation allows the evaluator precisely to assess how a student is interplaying parts so as to create meaningful wholes. This holistic complexity of thought and speech is exactly what standardized testing will never measure.

Standardized tests rank students or schools but don't troubleshoot the problems of either individuals or institutions. Language educators need to learn what to look for so as to become more expert assessors and counselors in the learning process. In organic,

student-centered language learning such as I have advocated elsewhere,* these two roles are synonymous. In order to guide students who are creating individual curricula while interacting with others, teachers have to evaluate constantly *and* have to teach students to do the same. All are charting past and future together on the basis of what previous activities have been worth and what is needed next. So both products and processes are examined all the time as members of reading and writing groups confer about texts, as writing is responded to and disseminated, as work folders fill up for perusal, and as texts are given rehearsed readings or otherwise performed live.

Discussions, improvisations, performances, writing workshops, and all sorts of projects can be frequently audiotaped and videotaped for several purposes at once—to allow participants to critique themselves, to furnish material for teacher in-service discussions, to orient new students to these activities, and to show people outside the classroom what is going on inside. Many of these tapes may serve only temporarily and then be recorded over. Others may be saved along with selections from writing portfolios and other tangible products to provide more lasting records. Doing both accommodates random, slice-of-life sampling of the whole curriculum and tracing of individual growth.

The point is to have plenty to look at so that any party can evaluate for any purpose. Observing processes as well as examining products permits realistic troubleshooting. Yet none of these activities exists only for assessment; they are all learning processes. They are the target activities, the goals themselves of speaking, reading, writing, and thinking—of communicating, collaborating, decision-making, problem-solving, creating, and interpreting texts—not some exercises that are means only, alleged to eventuate sometime in these goals. What you're seeing is what you're getting.

Whether external testing simply disappears in the future, or whether it does indeed become identical with the learning activities that are the goals, perceptive observation will emerge as the central means of evaluating. Assessment reformers today base their claim to be able to measure higher thinking on what they call the three Ps—performances, portfolios, and projects. If they succeed, they will

* See *Student-Centered Language Arts, K–12,* James Moffett and Betty Jane Wagner, 4th edition, 1992 Boynton/Cook, Portsmouth, NH.

The chapter in it titled "Evaluating" also elaborates some of the issues and processes touched on in this book. Consult the chapters there on talking, improvising, performing, reading, and writing for specific things to look for in each of these activities.

be assessing in the routine classroom or workplace, not in special examination circumstances on rare occasions under conditions that permit numerical comparisons. The three Ps sound very good, if they are authentic language activities as practiced out of school. Performances, portfolios, and projects make up the kind of curriculum this book presupposes. Made available in some slice-of-life form to outsiders, they offer real alternatives to the old multiple-choice tests for external examination. In any case, they all depend on observers knowing how to look and how to think about what they see.

Freed from emceeing to observe in their own classroom, teachers can note personal traits and trends, comment on these as needed during conferences with students and parents, and write reports if required. When students are doing different things according to personal experience and choice, they look different because they create individual patterns that are far more distinctive than test-score profiles. Teachers know more surely how to coach, confer, and counsel toward improvement. The descriptions of growth spelled out in this book aim to help teachers think about what to look for.

Some of the most important signs of verbal growth are certain habits, attitudes, and feelings too obvious to dwell on here but necessary to reaffirm. Well-developing language learners will feel more at ease speaking, reading, and writing and will consequently increase their fluency and pleasure in these areas. By exploring what oracy and literacy can do, students will increasingly appreciate the multiple uses of language—to socialize, play, communicate, think, and create. By finding out the limitations of language, they will discriminate between occasions for words and occasions for silence. Confidence and curiosity increase. Choices multiply. Expression acquires verve and subtlety; interpretation, justness. Thinking broadens and deepens. Let's never forget to look for and register major human developments in expression and understanding as we break these down now into more specific things to detect.