

Testing and Thoughtfulness

The literacy of thoughtfulness calls for a new concept of testing, one that reflects the active nature of learning.

The definition of literacy that sufficed for an earlier generation and a different economy has been replaced by a host of higher literacies: computer, scientific, civic, cultural, and so on. These higher literacies call for students to analyze; think critically, evaluate, synthesize information, communicate more effectively, solve problems, learn how to learn, and, in general, learn far more actively than traditionally. They call for greater thoughtfulness, both in the thinking sense and in the sense of caring more about other people and one's community.

For the past two years, with a grant from the John D. and Catherine T. MacArthur Foundation, we at the Education Commission of the States have been investigating the perceived and real effects of policy upon efforts to develop a much higher level of literacy for a much broader range of students. This "Policy and the Higher Literacies" project came about because education reformers, politicians, and business leaders are calling for larger numbers of graduates to be good thinkers, problem solvers, and risk takers.

Policy and the Higher Literacies

What we have been asking ourselves, are state and local education policies doing to promote greater thoughtfulness in school or to inhibit it? We have been looking carefully at policies in curriculum (including textbook selection), teacher certification and professional training, school finance, and testing and accountability. We have interviewed experts, held focus group meetings, reviewed all available research, and conducted case studies of districts in six regions.

In our case studies, we reviewed the

state and local policies and guidelines affecting thoughtfulness (e.g., mandates mentioning thinking or problem solving); interviewed appropriate state and local administrators; and then visited schools and classrooms (grades 3, 6, 8, and 11) where "higher literacy" activities were purportedly going on. All told, in these case studies we gathered 650 hours' worth of interview and observational information, about two-thirds of which was derived from talking with teachers and observing their classes.¹

Almost everyone we talked to is determining educational success or progress on the basis of scores on commercial standardized norm-referenced tests. This, of course, is not news to anyone who has been watching American education over the last 20 years. Schools, principals, and teachers differ markedly, however, in their perceptions about this fact and the significance they attach to it. What it means to use these tests depends on who you are and what kind of pressure you feel you are under. By and large, schools that are not in the public doghouse or under tight court monitoring do their testing routinely, make little use of the results, and go about their business. On the other hand, schools under public scrutiny (even for reasons having nothing to do with achievement) encourage much more teaching to the test and make much more out of declines or increases in scores.

Analogously, experienced, confident teachers pay little attention to externally imposed tests. They don't need them, except, perhaps, to reinforce judgments they have already made on the basis of their own more comprehensive information about students. If test results contradict what

they know about students, they tend to ignore them. New teachers, however, and underprepared or insecure or embattled teachers, are more likely to teach to the tests, even if they believe the tests are poor indicators of student learning.

Next, most people we talked to believe that commercial standardized tests are valid indicators of student and school achievement. Although many researchers and even testing experts see serious limitations to these tests and caution users constantly about misuse or misinterpretation, school people seldom showed awareness or concern about these matters. When we asked whether the tests adequately addressed higher-order thinking skills or performances (they do not), they either didn't know, didn't care, or believed the test manufacturer had dealt with that problem.

Third, most people we talked to about thinking, creativity, and problem-solving programs believed that these kinds of activities cannot be tested at all or cannot be tested objectively. They would invest in more innovative programs, they said, if inexpensive tests could be developed for them.

We also found that in certain large districts thinking skills are being taught by breaking thinking itself into a collection of facts, terms, and subactivities that can be taught and tested one at a time. Our observations of these approaches led us to conclude that whatever else was going on, students were *not* engaged in active thinking or problem solving or creativity. Thoughtfulness itself had been adapted to the prevailing basic skills transmission model of teaching and testing.

What we found, in short, was con-

Current and Future Possibilities in Testing

Innovation in testing and assessment is following four paths:

1. *Expanding existing tests and data-gathering instruments.* Pittsburgh, for example, has developed thinking skills tests for social studies and comprehensive essay examinations along European lines. States and districts are borrowing activities from the National Assessment of Educational Progress to assess higher-order thinking; are using multiple tests rather than relying on a single indicator of achievement; and are scoring essay examinations holistically or analytically. Efforts are under way to raise accreditation standards and give accreditation visits and reports the same clout and status test scores enjoy.

2. *Analyzing and packaging existing test information more innovatively.* Some schools are repackaging existing test information more attractively with graphics and marketing it more purposefully than in the past. Others are duplicating questions used on national education polls and distributing them locally. Still others are computing "satisfaction indexes" by analyzing various sources of information about the community; for example, data about parent organization participation and parent comments on back-to-school night.

3. *Adapting and legitimizing evaluation schemes and instruments used in other fields.* Educators are turning to tests of creativity, divergent thinking, and problem solving (e.g., Ennis-Weir Critical Thinking Essay, Cornell Critical Thinking Test), some of which have been normed, many of which are undergoing validation as they are being used. Writing samples that are part of existing tests are being examined for what they reveal about thinking and problem solving. The Education Commission of the States is working with schools that use the "teacher-as-researcher" concept to monitor and evaluate progress. Some schools are exploring the "walkabout" approach, one of several comprehensive "rites of passage" for students (e.g., exhibitions, performances, and portfolios). Clinical evaluation models, juries, and peer evaluation approaches can also be found.

4. *Breaking new ground.* Departures from traditional tests can be found in six areas:

- *Computer testing.* Adaptive computer tests lead students from one level of difficulty to another, building on what they know and teaching them as they test. Intelligent computer tutors do the same thing—fusing learning and evaluation and individualization.

- *Video evaluation.* The Key School in Indianapolis, Indiana, is experimenting with this high-tech approach to evaluate its unusual curriculum and pedagogy based on theories of multiple intelligences. A Virginia school uses the open access cable television studio to display student achievements to anyone who tunes in.

- *Tying evaluation to learning.* Debbie Meier's Central Park East school has been restructured around five fundamental questions that drive all activities in the school: How do I know what I know? What is the point of view behind that statement? How does this connect with anything else? What if . . . or Suppose that . . . ? Who cares, and what difference does it make? In such a school organized around questions, not answers, evaluation is inseparable from learning.

- *Student-created tests, assessments, evaluations, and research projects.* One 7th grade teacher I observed regularly requires his students to write test questions and critique tests as a natural part of the teaching process. For them, evaluation is no longer a mysterious process conducted by some external agent; it is something they can do themselves. They are internalizing standards and judging for themselves the quality of their work.

- *Climate assessment.* Examples in this area include the thoughtfulness inventory used in the Education Commission of the States study and literacy audits developed by the Center for Early Adolescence.

- *Input from the community.* Much as blue-ribbon panels often evaluate aspects of city government and other public institutions, schools are impaneling business leaders, politicians, and parents to judge the effectiveness of schools.

With all of these possibilities, no one should claim that robust outcomes "cannot" be assessed.

—Rexford Brown

could not be tested at all or at least not within the current accountability framework. Educators trying nominally to introduce critical and creative thinking and problem-solving activities had reshaped those activities so they would fit the prevailing technology.

The Active Nature of Learning

We did see thoughtfulness in action in several places. Ironically—and instructively—the most exciting programs were in a district that uses *no* standardized tests or assessments. No one there felt much need for them, either to prod the schools toward higher achievement or to measure current achievement. Yet the schools were among the very best we visited, and the students were among the most energized we met. We did not doubt that they were learning constantly and well, regardless of their backgrounds, which were as diverse as you can find.

The primary rub between thoughtfulness and the kinds of testing to which we are addicted is the rub between active and passive learning. The new literacy of thoughtfulness calls for a quite different technology of teaching and testing; it rests on quite different notions about the nature of knowledge and learning because it is about the making of meaning, not just the receiving of it. Thoughtfulness is a constructive, not a passive, undertaking.

The prevailing teaching and testing technology rests on the assumption that knowledge is objective and can be drilled into passive "blank-slate" brains, then paraded out on cue. This may apply to some kinds of knowledge—for instance, multiplication tables. But most knowledge is socially and personally derived, we are creating it all the time through social interaction, and its nature and uses constantly shift. Moreover, adults spend most of their time sifting through and using various "knowledges" rather than "arriving at" them. They have to

siderable passivity with respect to testing, mixed with lackadaisical use of results and a dependency that, under certain circumstances, leads to shaping of teaching and curriculum around the tests. People engaged in higher

literacy activities did not worry much about the standardized tests, correctly assuming their students would do fine on them. People who were looking for excuses not to undertake higher literacy activities complained that they

analyze, synthesize, interpret, and evaluate facts and ideas far more often than they have to "know" them in the sense of only being able to recite them. These are precisely the kinds of thoughtful abilities that more and more leaders say graduates should possess and that fewer and fewer graduates actually do possess. And they are precisely the activities that cannot be tested the way we currently test students in most schools.

A New Concept of Testing

What needs to change in order to make room for more thoughtfulness in the schools is not just the tests—which clearly do not require students to create, construct, negotiate, and communicate meaning—but the whole concept of testing itself. The concept that testing is initiated externally from the student, separate from the learning process, and primarily aimed at determining whether inert knowledge is in students' short-term memories exerts far too much influence over school people today. The goals of thoughtfulness are that students internalize capacities to evaluate their learning, do so as they learn, and do so in ways that exhibit their capacity to be *performing* thinkers, problem solvers, and inquirers. The dominant technology neither tells us whether those goals are being met, nor encourages anyone to find out, nor models a kind of inquiry into achievement that students and teachers might profitably imitate.

The best tests of thoughtfulness I saw in my travels were student generated. Nothing helps you master a subject better than having to ask and debate fundamental questions about what is most important about the subject and how you could tell if someone else had mastered it. After all, higher literacy is acquired more through a "do as I do" than a "do as I say" relationship between teachers and learners. It is developed through

experience more than through lecture, through practice more than through passivity.

Students of all ages who create some of their own examinations are forced to reflect on what they have studied and make judgments about it. They are forced to ask questions and challenge answers. They come face to face with the fact that 30 students derive 30 different interpretations from the same lesson, and they have to evaluate and reduce those differences. They learn how to ask the right questions and how to know good answers from bad, irrelevant from relevant. Students who have to perform or exhibit their knowledge and skills get learning in their bones: active learners become lifetime learners.

It may appear that if too many schools turned to open-ended testing, portfolios, performances, and exhibitions, no one would be able to compare their effectiveness, and policy de-

velopment would suffer as a result. This need not become a problem. There is no reason to do away entirely with standardized tests that do not promote cheating (e.g., everyone's being "above average"). The task is to keep them from dominating school affairs to such a degree that they distort curriculum and hold back efforts to make schools more thoughtful places. Moreover, professionals in other fields have found ways to aggregate "messy" information systems for policy purposes. This difficult but not impossible enterprise should have high priority in state departments of education and educational labs and centers.

Schools as Thoughtful Environments

What you as a professional can do right now is to scrutinize tests and testing practices and determine to what degree they constrain efforts to teach critical and creative thinking, problem solving, the reasoning arts, and other dimensions of thoughtfulness in your school. If you find evidence of constraints, start a sustained schoolwide inquiry into ways you can remove them and demonstrate learning in more concrete, robust, and exciting ways than filling in ovals. You may find that you will have begun restructuring your school without even trying. At the very least, you will create a thoughtful environment that will model the very behavior you most want to cultivate in your students. □

1. Our report, *Schools of Thought*, is in preparation.

Rexford Brown is Communications Director and Director of a project called "Policy and the Higher Literacies" at the Education Commission of the States (ECS) in Denver, Colorado. He also worked on the National Assessment of Educational Progress for 11 years when it was housed at ECS, developing assessments and interpreting the results.

Students who have to perform or exhibit their knowledge and skills get learning in their bones: active learners become lifetime learners.

Copyright © 1989 by the Association for Supervision and Curriculum Development. All rights reserved.