

# Teaching to the (Authentic) Test

Testing can once again serve teaching and learning if tests clarify and set intellectual standards.

Art courtesy Karen Bradford

Practical alternatives and sound arguments now exist to make testing once again serve teaching and learning. Ironically, we *should* "teach to the test." The catch is to design and then teach to *standard-setting* tests so that practicing for and taking the tests actually enhances rather than impedes education, and so that a criterion-referenced diploma makes externally mandated tests unobtrusive—even unnecessary.

## Setting Standards

If tests determine what teachers actually teach and what students will study for—and they do—then the road to reform is a straight but steep one: test those capacities and habits we think are essential, and test them in context. Make them replicate, within reason, the challenges at the heart of each academic discipline. Let them be authentic.

What are the actual performances that we want students to be good at, that represent model challenges? Design them by department, by school, and by district—and worry about a fair, efficient, and objective method of grading them as a *secondary* problem. Do we judge our students to be deficient in writing, speaking, listening,

Virgil describes Augustus in this manner in order to make him seem more powerful to the Roman citizens. He is blessed and revered in these passages as only a hero or great leader would be. The language Virgil uses here is similar to language used to describe gods in other epics, and it also puts Augustus on the level of Aeneas.

Many of the passages which refer to Aeneas can be shown to relate to Augustus as well. Both men are, in some way, "founding fathers." At the beginning of the epic, Virgil describes Aeneas as "A man apart, devoted to his mission" (I, 1.16, p. 3), and later he reflects: "destiny drove them on/ ...so hard and huge/ A task it was to found the Roman people" (I, 1.47, p. 4). Augustus is compared to Aeneas; he is re-founding the Roman People, and he, too, is devoted to his mission. Another comparison between the two leaders is made when Aeneas vows: "I shall not make Italians underlings/ To Trojans. For myself I ask no kingdom" (XII, 1.255, p. 374). Augustus several times refused to accept the title of Emperor, preferring instead to call himself princeps. "No and Aeneas <sup>was</sup> benevolent leaders who

*right //*

*actually, I think Virgil is more complimentary than he can be!*

*Virgil presents both Augustus and Aeneas as benevolent leaders who*

*Keep your focus on Virgil as primary source!*

--4--

artistic creation, research, thoughtful analysis, problem posing, and problem solving? Let the tests ask them to write, speak, listen, create, do original research, analyze, pose and solve problems.

Rather than seeing tests as after-the-fact devices for checking up on what students have learned, we should see them as instructional: the central vehicle for clarifying and setting intellectual standards. The recital, debate, play, or game (and the criteria by which they are judged)—the “performance”—is not a checkup, it is the heart of the matter; all coaches *happily* teach to it. We should design academic tests to be similarly standard setting, not merely standardized.

Reform of testing depends, however, on teachers’ recognizing that standardized testing evolved and proliferated because the school transcript became untrustworthy. An “A” in “English” means only that some adult thought the student’s work was excellent. Compared to what or whom? As determined by what criteria? In reference to what specific subject matter? The high school diploma, by remaining tied to no standard other than credit accrual and seat time, provides no useful information about what students have studied or what they can actually do with what was studied.

To regain control over both testing and instruction, schools need to rethink their diploma requirements and grades. They need a clear set of appropriate and objective criteria, enabling both students and outsiders to know what counts, what is essential—what a school’s standards really are. Until we specify what students must directly demonstrate to earn a diploma, they will continue to pass by meeting the de facto “standard” of being dutiful and persistent—irrespective of the quality of their work. And standardized testmakers will continue to succeed in hawking simplistic norm-referenced tests to districts and states resigned to using them for lack of a better accountability scheme.

### Exhibitions of Mastery

The diploma should be awarded upon a successful final demonstration of mastery

### The Rite of Passage Experience (R.O.P.E.) at Walden III, Racine, Wisconsin

All seniors must complete a portfolio, a study project on U.S. history, and 15 oral and written presentations before a R.O.P.E. committee composed of staff, students, and an outside adult. Nine of the presentations are based on the materials in the portfolio and the project; the remaining six are developed for presentation before the committee. All seniors must enroll in a yearlong course designed to help them meet these requirements.

The eight-part *portfolio*, developed in the first semester, is intended to be “a reflection and analysis of the senior’s own life and times.” The requirements include:

- a written autobiography,
- a reflection on work (including a resume),
- an essay on ethics,
- a written summary of coursework in science,
- an artistic product or a written report on art (including an essay on artistic standards used in judging artwork).

The *project* is a research paper on a topic of the student’s choosing in American history. The student is orally questioned on the paper in the presentations before the committee during the second semester.

The *presentations* include oral tests on the previous work, as well as six additional presentations on the essential subject areas and “personal proficiency” (life skills, setting and realizing personal goals, etc.). The presentations before the committee usually last an hour, with most students averaging about 6 separate appearances to complete all 15.

A diploma is awarded to those students passing 12 of the 15 presentations and meeting district requirements in math, government, reading, and English.

*Note:* This summary is paraphrased from both the R.O.P.E. Student Handbook and an earlier draft of Archbald and Newmann’s (1988) *Beyond Standardized Testing*.

Fig. 1. An Example of a Final Exhibition

for graduation—an “Exhibition” . . . As the diploma is awarded when earned, the school’s program proceeds with no strict age grading and with no system of “credits earned” by time spent in class. The emphasis is on the students’ demonstration that they can do important things.

—From the Prospectus of the Coalition of Essential Schools

The “exhibition of mastery,” proposed by Ted Sizer in *Horace’s Compromise* (1984) and a cornerstone of the “Essential School,” is one attempt to grapple with these issues. The intent of the exhibitions project is to help schools and districts design more authentic, engaging, revealing, and trustworthy “tests” of a student’s intellectual ability.

The reference to engagement is not incidental. The exhibition of mastery was initially proposed as an antidote to student passivity and boredom, not merely as a more valid form of assessment. The idea is to capture the interest value of an authentic test of one’s

**If tests determine what teachers actually teach and what students will study for—and they do—then test those capacities and habits we think are essential, and test them in context.**

ability, such as is often provided in schools by literary magazines, portfolios, recitals, games, or debates. Thus, "any exhibition of mastery should be the students' opportunity to show off what they know and are able to do rather than a trial by question . . ."<sup>1</sup>

The exhibition of mastery, as the name implies, is meant to be more than a better test. Like the thesis and oral exam in graduate school, it indicates whether a student has *earned* a diploma, is ready to leave high school.<sup>2</sup> The school is designed "backwards" around these standard-setting tests to ensure that teachers and students alike understand their obligations and how their own efforts fit in a larger context. Teachers "teach to the test" because the test is essential—and teacher designed.

But why institute a radically new form of assessment? Why not just improve conventional teaching and course-related tests? As the "Study of High Schools" documented, a major cause of the high school's inadequacies is the absence of direct teaching of the essential skills of inquiry and expression. Even in "demanding" schools, students often fail to learn how to learn. The culprit is discipline-based curriculums that lead to content-based teaching and testing: the essential (cross-disciplinary) habits and skills of reading, writing, questioning, speaking, and listening fall through the cracks of typical content-focused syllabi and course credits; as indicated, for example, when teachers say "I teach English, not reading."

A required final public exhibition of know-how ensures that those essentials are taught and learned. The final exit-level exhibition reveals whether a would-be graduate can demonstrate control over the skills of inquiry and expression and control over an intellectual topic that approximates the expert's ability to use knowledge effectively and imaginatively. A final exhibition provides students with an occasion to make clear, if only perhaps symbolically, that they are ready to graduate.

An exhibition challenges students to show off not merely their knowledge but their initiative; not merely their

**The high school diploma, by remaining tied to no standard other than credit accrual and seat time, provides no useful information about what students have studied or what they can actually do with what was studied.**

problem solving but their problem posing; not just their learning on cue, but their ability to judge and learn how to learn on an open-ended problem, often of their own design. The experience thus typically focuses on the essential skills of "inquiry and expression"—a synthesis that requires questioning, problem posing, problem solving, independent research,

the creation of a product or performance, and a *public* demonstration of mastery. Significantly, there is often a component calling for self-reflection and analysis of what one has undergone and learned.

Thus, a *final exhibition* is a misnomer in an important sense. Many Coalition schools provide a semester- or yearlong course, an adult adviser, and a committee to ensure that a student has adequate guidance, evaluation, and incentive (see fig. 1 for an example of a final exhibition from a Coalition school). The exhibition of mastery is as much a process as a final product, if not more so. The process of choosing topics, advisers, and committees and refining one's ideas and skills is a yearlong exercise in understanding and internalizing standards.

A similar approach to a diploma at the college level has been used successfully at Alverno College, Milwaukee, Wisconsin, for over a decade.<sup>3</sup> Assessment is a central experience, with coursework a means to a set of known ends: students must achieve mastery in the following eight general areas, with their progress in each area being charted on a multistaged scale:

1. effective communication ability,
2. analytic capability,
3. problem-solving ability,
4. valuing in a decision-making context,
5. effective social interaction,
6. taking responsibility for the global environment,
7. effective citizenship,
8. aesthetic responsiveness.

### **Performances: Better Classroom Tests**

Course-specific tests also have glaring weaknesses, not only because they are often too low level and content heavy. They are rarely designed to be authentic tests of intellectual ability; as with standardized tests, teacher-designed finals are usually intended to be quickly read and scored.

It seems wise, then, to talk about a move toward more intellectual performances in course-bound testing as a way of stressing the need to make tests more central, authentic, and engaging—as in the arts and athletics. (The

### An Oral History Project for 9th Graders

#### To the student:

You must complete an oral history based on interviews and written sources and then present your findings orally in class. The choice of subject matter is up to you. Some examples of possible topics include: your family, running a small business, substance abuse, a labor union, teenage parents, and recent immigrants.

Create three workable hypotheses based on your preliminary investigations and four questions you will ask to test out each hypothesis.

#### Criteria for Evaluation of Oral History Project

#### To the teacher:

- Did student investigate three hypotheses?
- Did student describe at least one change over time?
- Did student demonstrate that he or she had done background research?
- Were the four people selected for the interviews appropriate sources?
- Did student prepare at least four questions in advance, related to each hypothesis?
- Were those questions leading or biased?
- Were follow-up questions asked where possible, based on answers?
- Did student note important differences between "fact" and "opinion" in answers?
- Did student use evidence to prove the ultimate best hypothesis?
- Did student exhibit organization in writing and presentation to class?

*Note:* This example is courtesy of Albin Moser, Hope High School, Providence, Rhode Island. To obtain a thorough account of a performance-based history course, including the lessons used and pitfalls encountered, write to Dave Kobrin, Brown University, Education Department, Providence, RI 02912.

Fig. 2. An Example of a Test of Performance

term *exhibitions* would be reserved for those culminating graduation-level exercises designed to assess ability in the essentials underlying all coursework required for graduation.)

Designing performances implies a very different approach to standard setting than is implied by typical criterion-referenced tests or outcome-based views of mastery, though the instincts behind the designs are similar. Performances would ideally *embody* and *evoke* desired outcomes in authentic contexts. Too often, specifying only outcomes leads to tests that atomize and decontextualize knowledge: the testmaker designs a set of isolated pat exercises designed to elicit each desired outcome. Genuine tests of ability rarely provide such blatant cues and simple recall; they require us to have a repertoire, the judgment and skill to "put it all together" in one central challenge, repeatedly tried. (Imagine the assessment of music ability in a series of little exercises tried once, rather than through practice and performance of a complete piece in recitals.)

In sum, the goals behind the exhibition of mastery and the performance are to design standard-setting tests that provide more direct evidence of a student's intellectual ability; design tests that are thus able to stand by themselves as objective results; design more authentic intellectual challenges at the

**Reform of testing depends on teachers' recognizing that standardized testing evolved and proliferated because the school transcript became untrustworthy.**

heart of a discipline; and design tests that are more likely to engage students and motivate them to raise their own intellectual standards to do well on them. (See fig. 2 for an example of a performance that illustrates and illuminates these design standards.)

#### Toward More Authentic Tests

Exhibitions and performances sound fine on a schoolwide basis, you say, but districtwide or statewide? Isn't that too costly and cumbersome? I contend that the supposed impracticality and/or expense of designing such tests on a wide scale is a habit of thinking, not a fact. The United States is the only major country that relies so heavily on norm-referenced, short-answer tests instead of performance- and/or classroom-based assessment on a national level. In addition, a national committee on assessment in Great Britain has called for an exemplary system requiring flexible, criterion-referenced, and performance-based tests.<sup>4</sup> Many of the tests would be created by classroom teachers, who would be part of the standardizing process through "moderating" meetings to compare and balance results on their own and national tests.

In the U.S., more authentic skill assessment can now be found in various districts and states due, in part, to the work in writing assessment by the National Writing Project and its state offshoots (such as the California CAP writing test), and the American Council on the Teaching of Foreign Languages in the assessment of foreign language proficiency. Some states, such as Connecticut, have already designed and piloted performance-based assessment using ACTFL tests and criteria. In addition, they have piloted hands-on tests in graphics, small engines, and science. Vermont has proposed a statewide assessment system in writing and mathematics that would be portfolio based and teacher assessed.

We already have a national example in science: the 1987 NAEP pilot "Higher-Order Thinking Science Test," which includes some (though too few) hands-on experiments. One example:

Students are given a sample of three different materials and an open box. The

### A. Structure and Logistics

1. Are more appropriately public; involve an audience, a panel, and so on.
2. Do not rely on unrealistic and arbitrary time constraints.
3. Offer known, not secret, questions or tasks.
4. Are more like portfolios or a *season* of games (not one-shot).
5. Require some collaboration with others.
6. Recur—and are *worth* practicing for, rehearsing, and retaking.
7. Make assessment and feedback to students so central that school schedules, structures, and policies are modified to support them.

### B. Intellectual Design Features

1. Are "essential"—not needlessly intrusive, arbitrary, or contrived to "shake out" a grade.
2. Are "enabling"—constructed to point the student toward more sophisticated use of the skills or knowledge.
3. Are contextualized, complex intellectual challenges, not "atomized" tasks, corresponding to isolated "outcomes."
4. Involve the student's own research or use of knowledge, for which "content" is a means.
5. Assess student habits and repertoires, not mere recall or plug-in skills.
6. Are *representative* challenges—designed to emphasize *depth* more than *breadth*.
7. Are engaging and educational.
8. Involve somewhat ambiguous ("ill-structured") tasks or problems.

### C. Grading and Scoring Standards

1. Involve criteria that assess essentials, not easily counted (but relatively unimportant) errors.
2. Are not graded on a "curve" but in reference to performance standards (criterion-referenced, not norm-referenced).
3. Involve demystified criteria of success that appear to *students* as inherent in successful activity.
4. Make self-assessment a part of the assessment.
5. Use a multifaceted scoring system instead of one aggregate grade.
6. Exhibit harmony with shared schoolwide aims—a *standard*.

### D. Fairness and Equity

1. Ferret out and identify (perhaps hidden) strengths.
2. Strike a *constantly* examined balance between honoring achievement and native skill or fortunate prior training.
3. Minimize needless, unfair, and demoralizing comparisons.
4. Allow appropriate room for student learning styles, aptitudes, and interests.
5. Can be—should be—attempted by *all* students, with the test "scaffolded up," not "dumbed down," as necessary.
6. Reverse typical test-design procedures: they make "accountability" serve student learning (Attention is primarily paid to "face" and "ecological" validity of tests).<sup>1</sup>

1. Thanks to Ted Sizer, Art Powell, Fred Newmann, and Doug Archbald; and the work of Peter Elbow and Robert Glaser for some of these criteria. A more thorough account of them will appear in an upcoming issue of *Phi Delta Kappan* (in press).

Fig. 3. Characteristics of Authentic Tests

samples differ in size, shape, and weight. The students are asked to determine whether the box would weigh the most (and least) if it were completely filled with material A, B, or C. There are a variety of possible approaches . . . NAEP administrators used detailed checklists to record each student's procedures and strategies.<sup>3</sup>

NAEP borrowed most of its experiments from the British Assessment of Performance Unit tasks, which have been used (and reliably scored) in

Great Britain for a decade in reading, speaking, listening, math, and science.

Genuine tests *can* be widely implemented if we can overcome inertia and fatalism about current forms of standardized testing. Authentic, performance-based testing is a reality, not a romantic vision. There is also ample room for more intelligent design and use of conventional norm-referenced standardized tests.<sup>6</sup>

The state of Connecticut has developed a "Common Core of Learning," which lists objectives and criteria in all essential domains. Performance-based tests, built around criteria specified by experts in each field and involving tests administered by trained observers, are to be designed to honor those aims.

There are even standardized tests worth noting. ACT has developed a wide-ranging multimedia test of "general education knowledge and skills" called COMP, designed for colleges but easily adaptable to the high school level. The test uses art reproductions and audiotapes of news programs, for example, in testing writing and listening skills. On other items, students draft letters on different topics. There is even allowance for the student to respond orally on tape to a few test questions. The test takes six hours to administer, covers all the essential skills of inquiry and expression, and includes a 54-question self-assessment about one's patterns of activity related to each competency.

In sum, authentic tests have four basic characteristics in common. First, they are designed to be truly representative of performance in the field; only then are the problems of scoring reliability and logistics of testing considered. Second, far greater attention is paid to the teaching and learning of the *criteria* to be used in the assessment. Third, self-assessment plays a much greater role than in conventional testing.<sup>7</sup> And, fourth, the students are often expected to present their work and defend themselves publicly and orally to ensure that their apparent mastery is genuine. (See fig. 3 for a more thorough list of characteristics of authentic tests.)

### Toward a Performance-Based Diploma

The diploma by exhibition implies radically different standards for graduation. Instead of seat time or the mere accrual of Carnegie units, the diploma is performance based and criterion referenced. We may not be ready for the demise of age grading and social promotion; but if the harm done by standardized testing is to be undone, we need to redesign schools "back-

wards" around graduation-level standards of performance.

The performances and exhibitions should be designed prior to instruction, thus setting the school's standards in functional, not merely abstract and idealized, terms. Seeing them as add-ons to the traditional curriculum is to miss the point. How must the school be redesigned to support exhibitions or any form of exit-level standards? This should be the question behind "restructuring" and the source of vigorous debate among faculties and school board members. Design-

**An "A" in "English" means only that some adult thought the student's work was excellent. Compared to what or whom? As determined by what criteria? In reference to what specific subject matter?**

ing and institutionalizing exhibitions would better ensure, in other words, that the school had clear, coherent, and effective standards. Knowing the desired student abilities and work standards, as embodied in culminating performances and scoring criteria, would force key issues of policy: how will time, space, personnel, and other resources be best spent to ensure that diploma standards are met?

To talk with disdain of "teaching to the test" is to misunderstand how we learn. The test is the point of leverage—for learning and for reform. The issue is the integrity of the test: the genuineness, effectiveness, and apt-

ness of the challenge. The finals (and the criteria by which they are graded) set the standards of acceptable work in a course and a school—irrespective of noble language in school district reports or teacher intentions as reflected in syllabi. Legitimate and effective assessment is as simple(!) as ensuring that tests, grades, diploma requirements, and the structures and policies of the schools practice what we preach as essential. If we so honor our professed aims, the problems associated with standardized testing will take care of themselves.□

1. From *Horace's Compromise* (Sizer 1984), p. 68.

2. This (final) exhibition is patterned after the 18th century model of a public display of one's ability to engage in disputation: "... candidates for degrees expected to be academically tested at Commencement itself. Bachelor of Arts candidates prepared theses or topics on which they could be quizzed, and candidates for the Master of Arts submitted questions they were ready to defend. Titles of theses and questions were printed in advance and handed out at Commencement, and visitors often took the opportunity of challenging the candidates on their knowledge" (from the Harvard University Commencement program).

3. See the booklet *Assessment at Alverno College*, available from the college. For a history and an analysis of Alverno's program (as well as a general discussion of competency-based higher education), see *On Competency* (Grant, Elbow et al. 1979).

4. *National Curriculum. Task Group on Assessment and Testing: A Report*. Available from the Department of Education and Science (1988). This is a landmark document, outlining in readable prose a plan for intelligent and humane assessment.

5. From *Learning By Doing* (Educational Testing Service 1987).

6. See the excellent article by Dan Koretz of the RAND Corporation in the Summer 1988 issue of *American Educator*, which sums up the current controversy about norm-referenced state testing (the "Lake Wobegon effect" of each state being above average) and provides a useful set of guidelines for assessing assessment.

7. At Alverno, self-assessment is often the first level of proficiency. Thus, in the speaking requirement, students must give a five-minute videotaped talk—with the first evaluations given on the student's

self-assessment after watching the videotape.

#### References

- Alverno College Faculty. (1979/1985) *Assessment at Alverno College*. Rev. ed. Milwaukee, Wis.: Alverno College.
- Archbald, D., and F. Newmann. (1988). *Beyond Standardized Testing: Authentic Academic Achievement in the Secondary School*. Reston, Va.: NASSP Publications.
- Department of Education and Science and the Welsh Office. (1988). *National Curriculum. Task Group on Assessment and Testing: A Report*. London: Her Majesty's Stationery Office, Department of Education and Science, England and Wales. A brief "Digest for Schools" is also available.
- Educational Testing Service. (1987). *Learning By Doing: A Manual for Teaching and Assessing Higher-Order Thinking in Science and Mathematics*. A report on the NAEP pilot of performance-based assessment. A summary of the NAEP pilot of performance-based assessment. Princeton, N.J.: ETS. The full report: *A Pilot Study of Higher-Order Thinking Skills Assessment Techniques in Science and Mathematics*. ETS Report #17-HOS-80.
- Grant, G., P. Elbow, et al. (1979). *On Competence: A Critical Analysis of Competence-Based Reforms in Higher Education*. San Francisco: Jossey-Bass.
- Koretz, D. (Summer 1988). "Arriving in Lake Wobegon: Are Standardized Tests Exaggerating Achievement and Distorting Instruction?" *American Educator* 12, 2.
- Sizer, T. (1984). *Horace's Compromise: The Dilemma of the American High School*. Updated ed. Boston: Houghton-Mifflin.
- Wiggins, G. (In press). "A True Test: Toward Authentic and Equitable Forms of Assessment." *Phi Delta Kappan*.

#### Recommended Readings

- Alverno College Faculty. (1984). *Analysis and Communication at Alverno: An Approach to Critical Thinking*. Milwaukee, Wis.: Alverno College.
- Berk, R. A., ed. (1986). *Performance Assessment: Methods and Applications*. Baltimore, Md.: Johns Hopkins University Press.
- Bloom, B., G. Madaus, and J. T. Hastings. (1981). *Evaluation to Improve Learning*. New York: McGraw-Hill.
- Brooks, G. (1987). *Speaking and Listening: Assessment at Age 15*. Great Britain: The

# MADELINE HUNTER INSTITUTES WILLIAMSBURG, VIRGINIA July 8-16, 1989

THE PERFECT COMBINATION!!  
A Touch of History with a Twist of Hunter

DR. HUNTER RETURNS FOR THE  
SIXTH CONSECUTIVE SUMMER

---

**LEVEL ONE**  
—ESSENTIAL ELEMENTS OF INSTRUCTION—  
JULY 8-12

---

**LEVEL TWO**  
—CLINICAL SUPERVISION—  
JULY 13-16

---

VERY REASONABLE REGISTRATION RATES  
*Come and Enjoy* ●●●●●

- COLONIAL WILLIAMSBURG—
- YORKTOWN—
- JAMESTOWN—

"THE HISTORIC TRIANGLE"

#### INSTITUTE REGISTRATION INFORMATION:

Write to: Joseph V. Grebb  
Institute Director  
P.O. BOX 179  
Williamsburg, VA 23187-0179  
or Call: (804) 253-6783

#### SPONSORED BY:

**Williamsburg — James City County Public Schools**  
**JOHN E. ALLEN — Superintendent**

- Assessment of Performance Unit (APU), Department of Education and Science. APU material exists on the results of performance-based assessment in language, history, math, science, and history in primary and secondary schools.
- Elbow, P. (1986). "Trying to Teach While Thinking About the End" and "Evaluating Students More Accurately." In *Embracing Contraries: Explorations in Teaching and Learning*. New York: Oxford University Press. The former chapter originally published in Grant, Elbow, et al. (1979).
- Higgs, T., ed. (1984). *Teaching for Proficiency, the Organizing Principle*. Lincolnwood, Ill.: National Textbook Company and ACTFL.
- Sizer, T. (1986). "Changing Schools and Testing: An Uneasy Proposal." In *The Redesign of Testing for the 21st Century*. 1985 ETS Invitational Conference Proceedings. Princeton, NJ: ETS.
- Slavin, R., et al. (1986). *Using Student Team Learning*. 3rd ed. Baltimore: The Johns Hopkins Team Learning Project Press.
- Snow, R. (1988). "Progress in Measurement, Cognitive Science, and Technology That Can Change the Relation Between Instruction and Assessment." In *Assessment in the Service of Learning*. 1987 ETS Invitational Conference Proceedings. Princeton, NJ: ETS.
- Spandel, V. (1981). *Classroom Applications of Writing Assessment: A Teacher's Handbook*. Portland, Oreg.: Northwest Regional Educational Laboratory.
- Stiggins, R. (1987). "Design and Development of Performance Assessments." *Educational Measurement: Issues and Practices* 6, 3: 33-42. An Instructional Model (ITEMS), published by National Council on Measurement in Education (NCME). Comes with an Instructor's Guide.
- Stiggins, R. (January 1988). "Revitalizing Classroom Assessment." *Phi Delta Kappan* 69, 5.
- Wiggins, G. (Winter 1987). "Creating a Thought-Provoking Curriculum." *American Educator* 11, 4.
- Wiggins, G. (Winter 1988). "Rational Numbers: Scoring and Grading That Helps Rather than Hurts Learning." *American Educator* 12, 4.
- Grant Wiggins** is a special consultant to the National Center on Education and the Economy in Rochester, New York, and the Coalition of Essential Schools on issues of assessment and curriculum. He may be reached through the National Center.

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