

## State-by-State Comparisons Can Benefit Education

School administrators once called state-by-state comparisons "invidious." Now many of them impatiently await the data that may help them make more effective policy decisions.

In February and March 1990, the National Assessment of Educational Progress (NAEP) conducted its first trial round of state-by-state assessment. This maiden effort, open to as many jurisdictions as wished to participate, involves 37 states, two U.S. territories, and the District of Columbia. This year's survey, authorized by Congress in 1988, was limited to measuring the mathematical proficiency of 8th grade students in public schools. Meanwhile, NAEP's regular 1990 national assessment has been proceeding concurrently. This year it spans reading, math, and science in grades 4, 8, and 12, in both public and private schools.

NAEP is scheduled to conduct a second cycle of the "trial state assessment" in 1992. It will assess math skills

at both 4th and 8th grade levels and reading skills among 4th graders as well. It, too, will be voluntary, with participating states again bearing a portion of the costs.

Though modest in scale and experimental in nature, these two phases of the trial state assessment represent the first systematic and purposeful attempts to determine how the skills and knowledge of youngsters in individual states compare with those of their peers in other states and with the nation as a whole. If this experiment succeeds, and if Congress authorizes (and funds) the continuation and elaboration of state-by-state assessment as a regular component of future NAEP cycles, in time we can look forward to a steady flow of accurate, reliable, and timely data on our children's educa-

**We may disagree on the importance, validity, or priority of particular data, but who would argue that the schools and the public shouldn't know how their students are doing?**

**The message to policymakers is to hold schools accountable for achieving negotiated goals but not to legislate how they are to achieve such results.**

tional progress. And these data will be available to policymakers at the level of government the authors of the Tenth Amendment determined as the proper one to deliver educational services: the state level.

During the recent "excellence movement," it's no coincidence that the states have been the most vigorous policy units in their efforts to reform and renew American schools. Nor is it chance that state officials—governors, legislators, and chief state school officers in particular—have been the most ardent advocates of this NAEP expansion. These officials will be the primary consumers of the resulting data. The federal government has been entrusted with the responsibility to organize, manage, and shoulder most of the costs of this sizable effort to inform the states, and this is in keeping with Washington's oldest and most fundamental obligation in the field of education. Congress established the first federal education agency in 1866 "for the purpose of collecting such statistics and facts as shall show the condition and progress of education in the several states and territories." One hundred twenty-four years later, this mission endures.

The new NAEP endeavor is risky and controversial. All sorts of things could go awry during the fund raising, assessment, data analysis, and reporting stages of the project. Many educators are skittish, both about what the resulting information will show and about how it will be used. Aware of these apprehensions, Congress moved cautiously, enacting only a modest two-stage trial rather than concurring in the wholesale commitment to state-by-state assessment recommended by the executive branch. Congress will revisit the matter upon completion of the trials and an independent evaluation, to decide what the project's future will be.

**Some NAEP History**

Since its inception in 1969, NAEP has monitored trends in American elementary/secondary educational achievement in a number of subject areas. The resulting data have been illuminating and in some ways useful. But NAEP's architects had brilliantly—and intentionally—designed an assessment that could not be used to compare states, school districts, or schools, let alone individual students. In order to forestall such "invidious" comparisons, which at the time were fiercely resisted by school administrators, NAEP initially sampled by age (whereas educational practices are structured by grade) and reported its results item by



item (which were interesting only to curriculum experts). Data were published for regions of the country, as well as for the nation as a whole, but statistics on individual states were unavailable.

By design, in other words, NAEP yielded few results that were relevant to educational practitioners and policymakers. In fact, the underlying goal of the test was not to find out whether the 9-, 13-, and 17-year-olds being tested had learned what they had been taught in the classroom; it was to find out what they knew and could do, period.

In 1983, after many years of being administered by the Education Commission of the States, the NAEP grant

was won by the Educational Testing Service (ETS). ETS promised a "new design for a new era." A major objective of this new design was to make NAEP more pertinent to policymakers. And, pretty much as promised, ETS has provided features that made policymakers take notice:

- In addition to age sampling, grade level data were gathered. (The 1988 statute now mandates this.)

- Scale scores were used to summarize dominant features in the data, and units on the scales were made meaningful in relation to the content of the test—that is, questions were indicative of specified levels of difficulty—and

were written in terms that laymen could grasp.

- Reports were simplified, made more readable and attractive, and targeted to lay audiences.

- The NAEP reports were brought to the attention of the press in order to increase public exposure.

Yet, even in the practiced hands of ETS, all was not perfect. The data flowed slowly. The 1986 reading assessment was beset by a bizarre "anomaly"—a suspicious drop in the national reading performance trend, later found to be the result of minor changes in the test instrument—that rendered its data questionable. But the most important development, utterly



*Thirty-seven states, two U.S. territories, and the District of Columbia participated in this year's trial round of state-by-state assessment. This test, limited to 8th grade mathematics in public schools, is NAEP's first effort to provide comparisons of student achievement to state policymakers.*

*Here an assortment of students from Farragut Primary School, Knoxville, Tennessee—not to be interpreted as preparing for individual comparisons—symbolizes a cheerful and confident response to state-by-state comparisons.*

*Photograph by Klaus Studio, Knoxville*

unforeseeable when ETS sketched its "new design" in 1983, was the state-centered reform movement that suddenly eclipsed the national focus on education quality.

The obsession of the mid-1980s became student learning outcomes, but the data available to ardent state education reformers were wretchedly inadequate. So in 1984, Education Secretary T. H. Bell began publishing the celebrated "wall chart" in an effort to make such information available. This immense piece of paper displayed state-by-state comparative education statistics from teacher salaries to dropout rates. When it came to student achievement, however, the wall chart relied solely on average scores earned by college-bound high school seniors on the Scholastic Aptitude Test (SAT) and the American College Test (ACT).

Secretary Bell, acknowledging that these were shaky indicators of cognitive outcomes, pointed out that they were the best, and really the only, available data. Until something better came along, he said, they would continue to be used.

At the same time, a historic shift was taking place among the very groups that had been most wary of comparisons when NAEP was created. This could best be seen in the Council of Chief State School Officers' (CCSSO) 1984 agreement to develop a system of comparative state-by-state education indicators that would include assessment of student learning outcomes. In 1985, the "Chiefs" voted to endorse the expansion of NAEP as one mechanism for collecting those data.

In 1986, the National Governors Association, led by Tennessee's Lamar Alexander, issued its education report, *Time for Results*. In the chair's summary, Alexander wrote, "The nation—and the states and school districts—need better report cards about results, about what students know and can do." And the following year, when the association issued its first follow-up report, Chair Tom Kean of New Jersey was even more explicit: "The data we now have on education results are not good enough. We know about the inputs but not enough about the outputs. . . . [The governors] should sup-

**Finding ways to make fair comparisons when the number of states is few but the differences among them and their students are many will continue to be the greatest challenge to the trial assessment.**

port current efforts to expand the National Assessment of Educational Progress to measure these skills [and should] support comparisons among states."

In the meantime, ETS was independently responding to various requests from states to "piggy-back" state-specific testing onto the national assessment program. Wyoming and Georgia contracted to conduct concurrent state-NAEP assessments in 1986, and a number of southern states also conducted special concurrent state-NAEP assessments in a project organized by the Southern Regional Education Board during 1985-1987.

In 1986, Education Secretary William J. Bennett formed a prestigious 22-member panel to consider recommendations on how national assessment could be improved. He appointed Governor Alexander to chair the blue-ribbon committee and the

distinguished educator H. Thomas James to serve as vice-chair and study director. After a year of deliberations, the Alexander-James panel delivered its report in March 1987. Entitled "The Nation's Report Card," it recommended that NAEP begin to systematically provide state-by-state comparisons.

With the approval of the Reagan administration, a bill was sent to Congress that included this directive as part of its reauthorization of NAEP. It should be noted that nothing in the old law barred state-by-state NAEP and the executive branch might have proceeded unilaterally. But so momentous was this decision for the country—and so costly would it prove in terms of budget and appropriations for NAEP—that the administration decided to seek the agreement of Congress.

Anticipating legislative authorization for state assessment, the National Center for Education Statistics (NCES) joined with the National Science Foundation in 1987 to issue a grant authorizing the CCSSO to develop a set of objectives for the 1990 NAEP math assessment. The grant stipulated that the Chiefs must employ a consensus process to identify test objectives the states could agree on. With representatives from 18 national associations and curriculum subcommittees in each participating state enlisted as resources, the CCSSO embarked on this unprecedented venture. Its test objectives were later adopted by the Educational Testing Service and by the NCES, and they formed the basis for the development of the 1990 math assessment.

In 1988, Congress ratified the two-state trial state assessment (and many other significant changes in NAEP and in federal education statistics-gathering) as part of the Hawkins-Stafford Elementary and Secondary Education Improvement Amendments. The Senate readily embraced the concept of state-by-state assessment, but the House was more wary. It set limitations and checks on the process and confined it to the "trial" mode. A fundamental alteration made in NAEP by this landmark legislation was the creation of the independent National Assessment Governing Board.

This board was charged with formulating policy guidelines for NAEP (including such sensitive matters as what and how to test and how to report the data) and advising the NCES Commissioner on NAEP's conduct.

In 1988 these changes appeared ambitious, even audacious. Yet, barely a year and a half later, President Bush and the governors would emerge from their education summit in Charlottesville, Virginia, with a set of plans and commitments that would in all likelihood necessitate further large changes. The summit participants agreed to set national education goals and standards for the first time in American history and also committed themselves to produce annual "report cards" that will monitor and apprise the public of progress toward these goals and standards.

The President and the governors recognized that the states must be the primary recipients of such information. But, at the same time, this infor-

mation feedback system must also furnish its data so that "the information" can be compared, aggregated, analyzed, and reported across states and for the country as a whole. In short, we will need a system that furnishes more information than we've ever had before—one that lends itself both to the policy needs of individual states and to those of the nation as a whole.

Now, seven years after *A Nation at Risk* helped galvanize states to reform and strengthen their education systems, the first state-by-state assessment of educational progress has actually been conducted. If all goes well, the resulting data should be released in 1991. However, many issues must be settled before this occurs. Before discussing some of these, it is worth pausing a bit more over the question of just *why* state-by-state assessment is a good thing for American education. The mere existence of an appetite for data among policymakers does not itself prove that what they crave is desirable. Some cravings are worth resisting.

#### **State Comparisons: Their Value**

The public today gets confusing and conflicting information about how their state and local education systems are doing. These systems often give us a cheery picture of steady progress and above-average pupil attainment. That picture can get very rosy: all 50 states have managed to report that they are above the national average in elementary school achievement.<sup>1</sup> This "Lake Wobegon" view of our educational efficacy contrasts sharply with depressing national and international reports of young Americans' educational achievement. Recent NAEP results in reading, mathematics, science, and writing reveal a very different achievement picture for the nation as a whole:

The profile of what America's high school students appear to know and are able to do is particularly disturbing. Sixty-one percent of the 17-year-old students could not read or understand relatively complicated material, such as that typically presented at the high school level. Nearly one-half appear to have limited mathematics skills and

abilities that go little beyond adding, subtracting, and multiplying with whole numbers. More than one-half could not evaluate the procedures or results of a scientific study, and few included enough information in their written pieces to communicate their ideas effectively. Additionally, assessment results in other curriculum areas indicate that high school juniors have little sense of historical chronology, have not read much literature, and tend to be unfamiliar with the uses and potential applications of computers.<sup>2</sup>

To place these outcomes in a broader context, a recent international comparative study of six countries (a study using NAEP test items) found American students ranking last in mathematics and among the lowest-achieving in science.<sup>3</sup>

The public, subscribing to the view that educators and school systems should be accountable for what students actually learn, not just for teaching them, wants reliable information on the actual performance of its education system. And the public is not squeamish about comparisons in this domain: A full 70 percent of those responding to the 1987 Gallup education poll (and 75 percent of those with children in public schools) were in favor of state-by-state and school-by-school comparisons of achievement test results. Only 14 percent were opposed.<sup>4</sup>

Although it is not a comprehensive indicator system, the NAEP trial state assessment will provide new measures of the education performance of the states in relation to one another and to the country as a whole. Most people recognize that present tests of learning outcomes are not adequate for this purpose. SAT and ACT scores are not based on representative samples of students, and commercial tests provide misleading results when used for this purpose. Scores based on a test developed in one state cannot be compared to scores from tests developed in other states.

State-by-state comparisons of cognitive learning promise to sustain interest in education improvement, even as they monitor the progress of recent reforms. Not only will they illuminate whether some states are doing better than others, they will also help educators and policymakers diagnose with

**The mere existence of an appetite for data among policymakers does not itself prove that what they crave is desirable. Some cravings are worth resisting.**

greater precision which aspects of their education systems are vibrant and which are in need of therapy.

Policymakers can now look forward to a time when it will be possible not only to know how the students in one's state compare in overall terms to those in neighboring jurisdictions—and to the nation as a whole—but also to see whether science achievement, for example, is in better or worse shape than math or reading, whether strong results in the earlier grades are dissipated (or weak ones reversed) in the upper grades, whether some parts of a subject (e.g., expressive writing) are markedly ahead of others (such as persuasive writing).

This information is valuable far beyond the state's own boundaries. Several nationwide reform-and-renewal projects undertaken by policymakers—notably the National Governors Association and the Council of Chief State School Officers—cannot determine the progress they are making until interstate achievement data become available. The governors and the Chiefs have taken to leaving conspicuous "blanks" in their annual reports to show where they will put the NAEP data once this information is available. Perhaps the clearest illustration of their enthusiasm for the process is the larger-than-expected number of states and territories that volunteered to participate in the 1990 cycle—and to bear a portion of the costs of so doing. There is every reason to expect the number of participating jurisdictions to rise in 1992.

It must be noted, however, that limitations built into the current NAEP statute will keep policymakers from obtaining some of the information they now want. As late as 1993, when data from the 1992 assessment cycle should be available, no state-level information about learning outcomes in science, writing, geography, and history will be included. Information at the 12th grade level in every subject will also be absent. A further limitation, said to have been inserted into the legislation at the behest of local school boards and parent-teacher organizations, forbids the use of NAEP's state-by-state test items for generating

## **The "Lake Wobegon" view of our educational efficacy contrasts sharply with depressing national and international reports of young Americans' educational achievement.**

information on student achievement at the district, building, or pupil level. The state is the smallest education policy-and-delivery unit that NAEP is presently allowed to test. This means that the governors of Texas and North Carolina will be able to compare statewide student performance with that of other states but won't be able to use NAEP to compare Austin with Charlotte or even with Dallas. This, in turn, may increase the total test burden in states that want intrastate achievement data but are obliged to keep using other instruments to obtain them. (And unless elaborate "equating" procedures are followed, they won't be able to appraise the results of their state testing programs on NAEP scales, either.)

### **What the Trial Will Tell Us**

In spite of these limitations, the 1990 state-by-state NAEP will yield a great deal of useful information, and the 1992 cycle can be expected to provide still more.

The National Assessment Governing Board is still working on the precise analyses and reporting arrangements by which the trial assessment data will be examined and made public. In the meantime, the test development and assessment planning activities have been guided by several fundamental principles, mostly arising from the Chiefs' consensus planning process. These specify that the trial results should:

- represent a reasonable nationwide consensus about what is important for students to learn;
- be based on sound testing and psychometric practices;
- be obtained with a minimum of intrusion into instructional time;
- take into account different circumstances and needs that states face;
- include data on features of the school system that can be improved by policymakers and educators.

In addition to these guidelines, the Chiefs recognized that, within the real constraints of time and money, the best assessments use alternate modes of measurement. These assessment procedures should address the challenges that students face in the "real world" environment of schools and classrooms, as well as the hopes and expectations of experts concerning what they should be learning.

The 1990 state assessment, as noted, was limited to 8th grade mathematics. The actual test consisted of approximately 142 items covering:

- numbers and operations (30 percent of the total test);
- measurement (15 percent);
- geometry (20 percent);
- data analysis, statistics, and probability (15 percent);
- algebra and functions (20 percent).

Furthermore, 30 percent of the items required procedural knowledge, 40 percent tested conceptual understanding, and 30 percent were oriented to problem solving. Approximately 29 percent of the assessment consisted of open-ended items, and another 29 percent required the use of calculators. In each state, the test was administered to roughly 2,000 students attending 100 schools. Each student was tested for just one hour.

Cognitive achievement items are not the whole story, either. Consistent with the statutory requirement that data be collected by NAEP only if "directly related to the appraisal of educational performance," the assessment has gained enormously in analytic potential from information derived from background questionnaires. For students, this included indicators of socioeconomic status, the home learning situa-

tion (homework, television-watching, access to reading materials), instructional experiences in math, time spent studying the subject, and attitudes toward it. From the teacher questionnaire we will learn about instructors' training and experience, their classroom strategies in math, the materials that they use, and the extent of their involvement in curricular decisions. Questionnaires administered to principals will yield data on the demographic composition of the student body, school policies and facilities, the principal's own background and use of time, and other areas of inquiry. An "opportunity to learn" survey meant to illuminate the exposure that students actually had in school to the elements of math probed by the assessment was planned for 1990. It had to be jettisoned because the field test indicated that the instrument was not reliable, but a different version of this survey should be up and running in time for the 1992 cycle.

The trial state assessment data will permit state-by-state and state-to-nation comparisons on achievement and a host of policy-related issues. These include variations in course-taking patterns, equity issues associated with the distribution of educational resources, variations in instructional practices as these affect math achievement, relationships between teacher characteristics and student attainment, correlations between the principal's characteristics and pupil achievement, and the portion of variation in student learning that may be accounted for by differences among schools.

#### Some Cautions

During the wait for these important data, we must guard against technical problems and obvious mistakes, and avoid raising expectations that even an expanded NAEP cannot meet. Assuming for the moment that the various challenges of test administration can be met despite the multitude of test sites and the discrepant policies that states follow, there are important limitations we need to bear in mind when the data are reported:

- We cannot generalize to popula-

tions not represented in the sample. The target population in 1990 is 8th grade students in public schools. This means that ungraded and nonpublic schools are excluded from the sampling frame.

- We cannot generalize to populations of students for which the sample was not designed to produce estimates. The sample contains some students who are handicapped, limited in their English proficiency, enrolled in vocational education programs, and receiving Chapter 1 services. Yet the state samples are not designed to be representative of these groups. Hence state estimates cannot be supplied for the achievement of these subpopulations. (In most cases, the only student groupings for which we will have reliable intrastate data in 1990 are by gender, race/ethnicity, and rural/urban residence.)

- We cannot generalize to variables not represented in the study. The trial state assessment will report only on what 8th grade students know and can do in mathematics. The report will not constitute a comprehensive description of student performance in the state's education system as a whole.

- We mustn't exaggerate what we know about pupil attainment, even in 8th grade math. Test scores are fallible proxies for achievement. A different

test, even in the same subject and grade, will give a different ranking of the states. Furthermore, some of what is tested has not been emphasized in the curriculums of all states, while some of what is stressed in particular states will not be on the test. A comprehensive appraisal of math achievement in the 8th grade would have to include much more than is planned in the trial assessment.

It is also important for the public to understand that the relative standings of the states are influenced by factors other than what is taught and learned in school. For example:

- States differ in their wealth and poverty, in the racial, ethnic, and socioeconomic composition of their populations, in the content of their curriculums, in their fiscal commitment to education, in their proportions of private (and ungraded) school enrollments, and in their definitions and percentages of handicapped and limited-English-proficient students.

- Variations in promotion and retention policies will yield state differences in the age distribution of 8th grade students.

- States differ in their percentages of migrant students, who are not good barometers of the educational outcomes of the state in which they happen to reside at the time of the assessment.

Finding ways to make fair comparisons when the number of states is few but the differences among them and their students are many will continue to be the greatest challenge to the trial assessment. Making wise policy in this domain will be perhaps the weightiest burden of the National Assessment Governing Board over the next several years, and the board has set up various advisory and consultative mechanisms to assist it in the quest for wisdom and fairness. But there is no reason to be paralyzed by this concern. Because it is so difficult to figure out how to produce analyses of student performance that are "fair," states themselves commonly disregard demographic differences when comparing students, schools, and districts on extant state achievement tests.

Statistical land-mines lie ahead, too. We need, for example, to guard

**NAEP makes no rules for American education. It hires no teachers, runs no schools, and confers no diplomas. NAEP merely tells the nation how its education system is doing.**

against the "ecological fallacy" of making inferences to individuals from aggregate data. We know that regression parameters obtained from means are not necessarily the same as those obtained from individual observations. This will be a particular problem when shifting the unit of analysis from the *individual student* in the intrastate NAEP reports to the *state* in the interstate report.

### **The Effect on the Measurement of National Trends**

Important though state assessment is, NAEP's top priority—plainly spelled out by the Congress, the Alexander-James panel, and the National Assessment Governing Board—remains the reliable measurement of national trends. The state-by-state program must not be allowed to disrupt or distort that essential task. But this is not easy, for the 1988 legislation requires NAEP in 1990 and 1992 to conduct two partially competing programs. One will continue to monitor national trends in student achievement. The other is a state testing program that, over time, could have "high stakes" for particular jurisdictions. As stakes rise, people become more apt to seek ways to beat the system. If states do this by altering their curriculums or teaching to the test, we may have created a testing environment that did not exist before. In other words, we "may have succeeded only in corrupting the inferences" we wanted to make from the national test.

Plans for the national and state assessments have features designed to diminish (though they cannot eradicate) this problem. For example, the test objectives are somewhat different. The national trend data for 1990 will be based on objectives developed for the 1986 NAEP cycle, while the test used in the trial state assessment will be based on the objectives developed through the CCSSO planning project.

The mode of test administration also differs. The national math trend data will come from a paced-and-taped administration, as has been done in the past. The state assessment relies on a timed-but-printed test administration.

## **Perhaps the most vexing instructional issue embedded in state assessment (as in any "high stakes" accountability program) is "teaching to the test."**

The data will be collected in different schools at different points in the calendar. The national trend information (for 13-year-olds) was collected in the autumn, whereas the state assessment (for grade 8) has been conducted in the late winter.

Precautions such as these reduce the likelihood that the state assessment program will distort the national trend data. But it should be noted that the added complexity of this two-track arrangement—juggling the competing demands of monitoring long-term national trends while beginning a new effort to compare states—also boosts the costs and the complexity of NAEP as a whole.

### **Will Curriculum and Instruction Be Affected?**

Congress has given the National Assessment Governing Board the task of "identifying appropriate achievement goals for each age and grade in each subject area to be tested under the National Assessment." The board has been considering several approaches to this portentous responsibility but at this writing had not settled on particular strategies.

The implication of this statutory provision is, of course, that NAEP will acquire a "normative" rather than merely descriptive character, not reg-

ulating anyone's behavior but making clear what it—or at least its governing board—thinks that behavior *ought* to result in by way of student achievement. If those goals are taken seriously by educators and policymakers, and if state-by-state assessment gauges and reports the performance of particular jurisdictions in relation to them, it is reasonable to expect that over time those goals and the test objectives derived from them will become factors in state and local curriculum decisions.

State assessment will almost certainly remain voluntary, however, and educators and policymakers will continue to be free to disregard some or all of it. Moreover, there is an important distinction between identifying appropriate goals and suggesting the curricular and pedagogical arrangements by which these may be attained. Broad learning outcomes can be set forth without prescribing the curricular scope-and-sequence, specific course offerings, modes of instruction, grouping practices, or the choice of textbooks and other class materials.

The objectives of the trial state assessment for 1990 were established (and those for 1992 are similarly being developed) through an extensive bottom-up consensus process that incorporated input from participating states. The objectives for the 8th grade math assessment are extremely broad and forward-looking.

Goals and objectives are not, in any case, the same as pedagogy. Similar educational outcomes can be obtained through many different classroom strategies—a good thing, given what is known about the variegated learning styles (and rates) of students and the diverse enthusiasms and competencies of teachers.

Perhaps the most vexing instructional issue embedded in state assessment (as in any "high stakes" accountability program) is "teaching to the test." This practice can have pernicious effects, such as restricting the content that students are taught in order to focus their attention on test preparation and eroding the concurrent validity of the test as a measure of achievement. In other words, teaching to the test reduces the correlation of

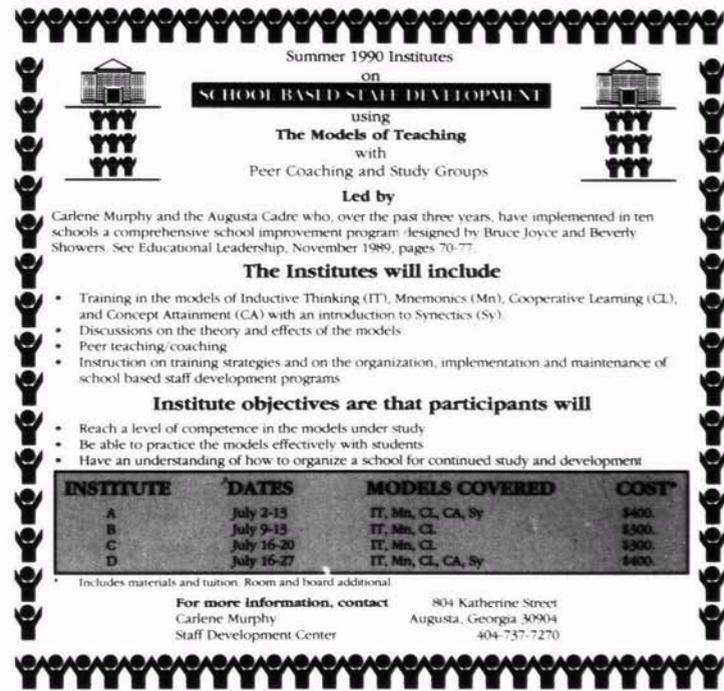
the test scores with alternative measures of the same domain.

As noted by Mehrens and Kaminski,<sup>6</sup> there is an ethical/unethical continuum in teaching to the test. The ill-defined point where one crosses over from legitimate to illegitimate practice depends on the inferences one wants to make from the test scores. But if one is satisfied that the test probes pretty much everything that is important for students to learn about the subject at hand, some degree of teaching to the test is appropriate, even commendable.

NAEP's goal must be to recommend national achievement targets and to provide a rich supply of information about state and national educational attainment, while preserving state and local control of curriculum and instruction. This is a tricky endeavor, which may be why Congress was ambiguous about it, leaving it largely to the National Assessment Governing Board to resolve.

Yet the board functions "in context," too. It is not the only entity suggesting achievement goals for American education, especially in the wake of the Charlottesville summit, and it must heed and try to cooperate with the others, which now span the National Council of Teachers of Mathematics, the American Association for the Advancement of Science, the Council of Chief State School Officers, the National Governors' Association, and the President's Education Policy Advisory Committee. If NAEP diverges sharply from the nascent (and ever-fluid) national consensus about what young Americans ought to know and be able to do, it will be rejected as an outdated source of meaningful indicators, much as a mechanism for gauging the performance of horses would be disregarded in the automotive age.

One of the virtues of the federal role in education is that it is auxiliary, not governing. Programs such as NAEP come into being as needs are felt and demands generated outside Washington, D.C. That is why national assessment began in the 1960s and why state-by-state NAEP is taking shape, albeit on a limited basis, in 1990. NAEP



Summer 1990 Institutes  
on  
**SCHOOL BASED STATE DEVELOPMENT**  
using  
**The Models of Teaching**  
with  
Peer Coaching and Study Groups  
Led by

Carlene Murphy and the Augusta Cadre who, over the past three years, have implemented in ten schools a comprehensive school improvement program designed by Bruce Joyce and Beverly Showers. See *Educational Leadership*, November 1989, pages 70-77.

**The Institutes will include**

- Training in the Models of Inductive Thinking (IT), Mnemonics (Mn), Cooperative Learning (CL), and Concept Attainment (CA) with an introduction to Synectics (Sy)
- Discussions on the theory and effects of the models
- Peer teaching/coaching
- Instruction on training strategies and on the organization, implementation and maintenance of school based staff development programs

**Institute objectives are that participants will**

- Reach a level of competence in the models under study
- Be able to practice the models effectively with students
- Have an understanding of how to organize a school for continued study and development

INSTITUTE	DATES	MODELS COVERED	COST*
A	July 2-13	IT, Mn, CL, CA, Sy	\$400
B	July 9-13	IT, Mn, CL	\$300
C	July 16-20	IT, Mn, CL	\$300
D	July 16-27	IT, Mn, CL, CA, Sy	\$400

\* Includes materials and tuition. Room and board additional.

For more information, contact  
Carlene Murphy  
Staff Development Center

804 Katherine Street  
Augusta, Georgia 30904  
404-737-7270

makes no rules for American education. It hires no teachers, runs no schools, and confers no diplomas. NAEP merely tells the nation how its education system is doing. Soon it will be able to provide some of that information to participating states as well. If the data prompt changes in policy and practice, so be it. If the information is entirely unhelpful or inappropriate, it is sure to be ignored. □

<sup>1</sup>J.J. Cannell, (1987), *Nationally Normed Elementary Achievement Testing in America's Public Schools: How All Fifty States are Above the National Average* (W.V.: Friends for Education).

<sup>2</sup>A.N. Applebee, J.A. Langer, and I.V.S. Mullis, (1989), *Crossroads in American Education* (Princeton, N.J.: Educational Testing Service).

<sup>3</sup>A.E. Lapointe, N.A. Mead, and G. W. Phillips, (1989), *A World of Differences: An International Assessment of Mathematics and Science* (Princeton, N.J.: Educational Testing Service).

<sup>4</sup>A.M. Gallup, and D.M. Clark, (September 1987), "The 19th Annual Gallup Poll of the Public's Attitudes Toward the Public Schools," *Pbi Delta Kappan* 69, 1:17-30.

<sup>5</sup>R. Brandt, (1989), "On Misuse of Testing: A Conversation With George Madaus," *Educational Leadership* 46, 7: 26.

<sup>6</sup>W.A. Mehrens, and J. Kaminski, (1989), "Methods for Improving Standardized Test Scores: Fruitful, Fruitless, or Fraudulent?" *Educational Measurement: Issues and Practice* 8, 1: 14-22.

*Authors' note* The views expressed in this article do not necessarily represent those of NCES, NAGB, or the Department of Education. Portions were presented by Phillips at the annual Assessment Conference sponsored by the Education Commission of the States in Boulder, Colorado, June 1989.

**Gary W. Phillips** is Chief of the Educational Assessment Branch at the National Center for Education Statistics, 555 New Jersey Ave. N.W., Room 310C, Washington, DC 20208. **Chester E. Finn, Jr.**, a former Assistant Secretary of Education, is Professor of Education and Public Policy at Vanderbilt University. He can be reached at 1112 16th St. N.W., Suite 500, Washington, DC 20036. He presently serves as chairman of the National Assessment Governing Board.

Copyright © 1990 by the Association for Supervision and Curriculum Development. All rights reserved.