We have witnessed a series of test-based educational reform efforts during the past two decades. In the 1970s it was the minimum competency testing movement. That movement had no sooner begun to lose steam when the use of tests as a major instrument of educational reform was given new impetus by the publication of A Nation at Risk and a rash of related reports on the status of education in 1983. Those reports used test results to make the case that reform was needed. More importantly, testing was stressed as a powerful reform tool.

New or expanded test requirements were a major component of the vast majority of reforms that were introduced throughout the 1980s. Greater emphasis on test-based accountability increased the stakes attached to test results, becoming a focus of intense debate. Several national organizations and commissions argued that there was an excessive amount of testing in the nation's schools. Not only was the efficacy of test-based reforms challenged, but the meaning of the test results was called into question by claims that the scores were inflated and provided misleading results about student achievement. Although numerous explanations could be advanced to explain the results, the Lake Wobegon effect (the tendency for all states and most districts to report results above the national average) undermined confidence in test scores.

Clearly, the idea of using tests as a major tool of educational reform is not new. The current movement to place greater reliance on testing has some new wrinkles, however. Today's proponents expect to overcome the shortcomings of previous efforts at test-based reform. Possibly the most important of these new wrinkles is the change in the conception of testing itself—a change considered dramatic enough to demand new terminology. Although the terminology is still somewhat in flux, (performance-based, authentic, or alternative) this new assessment labeling is intended to convey something different than the standardized tests which were relied upon so heavily for earlier test-based reforms.
essays, open-ended mathematics problems with multiple possible solution paths and sometimes with multiple correct answers, hands-on science tasks, and portfolios of work that may be accumulated over a year or more. These assessments are intended to be so consistent with instructional goals that teaching to them would not only be acceptable, but would be considered exemplary instructional practice. At the extreme, performance assessments become so integrated with instruction as to be virtually indistinguishable from instruction itself.

Evident from the increasing number of states that have moved aggressively to introduce performance-based assessments, state legislators have seized upon this vision of assessment as a potentially powerful agent of reform. Performance-based assessments have also been embraced to varying degrees by the National Education Goals Panel, the National Council on Education Standards and Testing, and the Department of Labor Secretary’s Commission on Achieving Necessary Skills. A key goal common to these recent efforts is that the assessments should encourage better instruction as well as greater student effort. America 2000, for example, calls for assessments that will foster good teaching and learning.”

It is easy to support the goal of fostering such important concepts. Assuring that the goal is actually achieved, however, is likely to be much more difficult. It will take more than rhetoric and good intentions if the present performance-based assessment approach to educational reform is to be more successful than the test-based reforms of the last two decades.

The development of effective and valid performance-based assessments requires considerable time, clear conceptual models, and careful empirical work. We fear that the present policy movement for such measures will short-cut the process and, in the long run, will result in assessments whose validity cannot be supported. Studies anchoring assessments in cognitive requirements of learning, instructional contexts, and construct validity must be a first priority if new assessments are to do a better job of living up to expectations set for them.
The Student Assessment Consortium is off and running," said Ed Roeber, Council of Chief State School Officers (CCSSO), “with strong state participation in six major topic areas. Our consortium in the fields of mathematics, reading, writing, the arts, social studies, and workplace readiness are busy meeting and developing a variety of new authentic assessments.”

Previously known as the Student Assessment Exchange, the Consortium is co-sponsored by CRESST and CCSSO—two groups committed to sharing the positive results of new standards and assessments. The Consortium held a major meeting on October 6, 1991 in Des Moines, Iowa.

According to Roeber, one of the greatest obstacles each workgroup has had to overcome is the diverse nature of state assessment programs. In mathematics, for example, some states have strict state assessment requirements that limit their flexibility in developing alternative assessments. Other states have few controls while still others have already made significant progress toward creating new forms of measurement.

Individual Consortium Actions
Each of the workgroups has met at least once, discussed issues, and established goals. For example, the reading and arts consortia are actively working on the development of shared standards. Comparing frameworks and existing assessments has been a priority for the social studies and workplace readiness workgroups.

“The social studies consortium wants to push for an integrated use of knowledge and skills across the social studies disciplines,” according to chair Pam Aschbacher. The workgroup also recommends that they address not only content knowledge but process and metacognitive skills and values.

Each consortium has suggested some innovative ideas for networking information between one another. For example, an electronic communication system to share outcome statements and performance items is a primary goal of the reading group. The writing prompt bank, developed by the National Writing Consortium, has been identified as a key existing resource for states interested in sharing writing assessments. Members of the writing group have also suggested the exchange of ideas across one or more of the other subject areas, such as social studies or reading.

The Future
Additional meetings have been planned by each consortium, generally at existing conferences where each participating state is already likely to be represented. The mathematics consortium will meet on April 1, 1992 at the National Council of Teachers of Mathematics (NCTM). The workplace readiness group will meet on February 18-19, 1992, in Washington, D.C., sponsored by the U.S. Department of Labor. They will meet again during April’s American Educational Research Association (AERA) conference in San Francisco. The social studies consortium will meet at CRESST/UCLA in March and has invited preeminent field experts to provide input.

“What we need now is funding for all of these groups,” said Ed Roeber, “and CCSSO is pursuing various options. I hope that something positive will happen in this area and soon.”

SCASS
A related assessment consortium has been formed under the guidance of CCSSO. Proposed by Thomas Boysen, Kentucky’s commissioner of education, the States’ Collaborative on Assessments and Standards for Students (SCASS) is surveying states to see what types of assessment exercises are needed and in which subject areas. SCASS represents those states financially able to develop new authentic assessments and standards immediately.

“The important thing is that we are finally moving forward in exchanging new standards and assessment information,” concluded Roeber, “and that the long-term benefits hold tremendous promise for our children.”

Further Information
Those states interested in further information should contact Ed Roeber at: CCSSO, Student Assessment Consortium, One Massachusetts Avenue, N.W., Suite 700, Washington, D.C. 20001-1431, (202) 408-5505. Other information should contact a consortium chair:


Reading: Barbara Kapinus, Maryland State Department of Education, (301) 333-2314.


The Arts: Joan Peterson, California Department of Education, (916) 657-4847.


Workplace Readiness: Stanley Rabinowitz, Far West Laboratory, (415) 565-3030.
Teachers As Researchers

by Charlotte Higuchi

During the 1991 CRESST conference, Charlotte Higuchi, a teacher from Farmdale Elementary School in East Los Angeles, made an impassioned plea about the need to recognize teachers as researchers, citing her own research into performance-based assessments. The following is an edited version of her thought-provoking and fervent comments which we think are important issues of interest to all education policymakers, researchers and practitioners.

Society holds up a standardized test and says to the teacher: “This is what we are going to test our children on and this is what you are going to teach.” I don’t do that in my class.

My children are all Chapter One students. My classroom is made up of bilingual, Hispanic, Vietnamese, and Chinese children, and one orthopedically handicapped child. They are all ungraded primary.

For me, learning is many pathways, not necessarily straight, but meandering, coming back, going forward. That’s what we don’t have in America. Right now the system is structured for the factory model. You have basal readers, a schedule you just put people through. The promise of alternative assessments is that now we’re going to look at each individual child and see what he or she can do.

Three years ago, the AFT (American Federation of Teachers) gave me a grant for $30,000 to develop and research performance-based assessments in my classroom—because I believe and I know that they will make a difference for our children. A friend of mine said, “feelings give intellect vitality,” and that’s what performance-based assessments do.

Alternative assessment is a natural part of the kind of curriculum changes that people are advocating today. The (performance-based) assessments that I use directly assess what my students can do. “Can they read? Can they write?” That seems simple but it’s not, because there are many things you have to consider for an open-ended kind of assessment. Authentic assessment helps me to know my children very well, and that kind of knowledge helps my instruction—I can change what I need for that child.

My children are now self-evaluating. Self-evaluation is a very rigorous intellectual process and it’s influenced by developmental stages. A teacher must know what is a “four-year old, five-year old, seven-year old,” not only intellectually, but emotionally, physically, personally. And teachers must know how all these attributes come together in one human being and how these attributes determine or color how our children are going to learn.

Authentic assessment cannot be done by just anybody. A substitute was covering my class yesterday and she said “if you are going to do these kinds of assessments, that means you have to change the way teachers are trained because they can’t do that now. And by the way, where are they going to get the materials to do this because you have to change the instruction?” The point is that teachers have to be trained in how to develop assessments and they have to have the materials to do it.

Unfortunately, teachers are still looked down upon. We are called stupid, uncaring and undedicated, and I resent it. Society has just about forgotten that teachers, who are with our kids so much of the time, know the kids best. If you read the great scientists like Einstein or Richard Feynman, you will find that what they say is this: “Science is based on knowledge, but it is also based on your imagination and intuition,” and that’s what an experienced teacher has. Unfortunately teacher knowledge and teacher research is not valued in this country.

Why do the Japanese and German students do better [than ours]? It’s not because they are smarter than we are, it’s because their society supports them. It’s not that Americans don’t value education, it’s just not their real priority.

This society puts teachers through a system in higher education that is totally irrelevant to what I do as a classroom teacher. Whatever I know today, I know because I’ve gone out and found out about it myself. I don’t use basal, I don’t use dittos or workbooks. All the work my children do is student generated. To do that kind of curriculum means you must have a vast amount of material. I have in my classroom 5,000 books, books that I bought myself or through the grants I’ve received. I’ve gotten all this [books and computers] not because society supports me, but because this is what it takes to teach.

Society is saying now, “oh this is wonderful, let’s do hands-on science, hands-on math, let’s have kids read books.” But where are the books? Where are the dictionaries? Where are the encyclopedias in the classroom?

Financial support makes a difference. My students’ parents care—last year they raised $2,000 so that at the end of spring I could take my children on an astronomy camping trip. That’s the kind of partnership we have. Black Elk, a Sioux holy man, once said “no one person can make a good thing happen.” What this implies for all of us is that we must work together to reform education.

(continued on page 7)
Task Force Recommends New Assessment System

The following article has been edited from the Assessment Task Force report to the National Council on Education Standards and Testing. The task force, chaired by Eva Baker, was comprised of distinguished education researchers from across the country. (Please see listing at end.) We anticipate that a complete report, including papers from the various task force members, will be available from CRESST in the near future.

Our Vision: Toward a System to Assess the National Standards

The Assessment Task Force recommends to the National Council a fundamental redesign of our assessment system—a redesign that can start immediately but will take time and care to come to fruition.

- We support an assessment system that protects children from the harms of test misuse, from unfairness, and from poor quality tests and assessment.
- We support an assessment system that holds schools accountable for providing high quality educational programs and for producing the intellectual accomplishments Americans expect of all of their children.
- We support a system that fosters rather than inhibits creative and demanding teaching.
- We need an assessment system to help us assess our progress toward the National Education Goals, and one that provides understandable information that will help mobilize families, students, educators, and the whole community to rededicate themselves to the cause of learning.

Overall Requirements for an Assessment System to Assess the National Standards

The new assessment system must be designed to ensure that states and local districts have the primary responsibilities for creating and implementing assessments for the purposes of accountability, school evaluation, student certification, reporting to parents, and instructional improvement. Such responsibilities should be State and local because decisions about schooling are made primarily at these levels.

The responsibility for assuring the quality of these assessments should reside at the national level. The national level also should be responsible for monitoring progress toward the National Goals (this responsibility could be fulfilled by using NAEP, as it is modified to reflect emerging national content standards). Additional national responsibilities are independently evaluating the impact of the assessment system on the Nation’s educational quality and equity, and providing resources for technical assistance, research, and development on assessment so that the system may be improved.

Assuring Quality of Assessments

Assessments must be judged to be consistent with the National Standards and must meet the criteria of validity, reliability, and fairness. Quality standards for the development and use of assessments should be developed and adopted by a nationally authorized entity. The membership of this entity should have the technical background to fit with their charge: experts in measurement, assessment, subject matter, learning, classroom processes, and special populations.

These quality standards should be adapted as appropriate from the Standards for Educational and Psychological Testing (AERA, APA, NCME, 1985) and should also include the eight criteria articulated in "Criteria for Evaluation of Student Assessment Systems" (National Forum on Assessment, 1991). Such quality standards, incorporated in promulgated guidelines, also should be helpful to states and local districts in the development of assessments.

Standards of Assessment Validity

Validity requires that the assessments be aligned with the national content standards proposed for development. The entity should assure that the assessments proposed by states or groups of states are aligned with the national content standards.

Furthermore, the entity should create guidelines and conduct studies to determine the comparability of assessments from different states or groups of states. At the outset, guidelines should be promulgated and reviews should relate to the comparability of assessment design in content quality, challenge to students, and content coverage.

Standards of Assessment Equity

The entity also should be responsible for devising criteria to assure equity of assessment. Special procedures will need to be developed to take language and culture into consideration for appropriate assessment. States should come forward with their plans for ensuring equity in assessment design, administration, and use for review by the entity.

Monitoring System Progress Toward Goal 3 of the National Education Goals

The most effective and efficient method of monitoring progress toward Goal 3 of the National Goals is through the National Assessment of Educational Progress (NAEP). NAEP, with the oversight of the National Assessment Governing Board (NAGB), should continue in its present role. Although development and refinement is necessary for NAEP to assess the content standards, NAEP’s existing design is well suited to meet this monitoring function for the nation and the states.

(continued on page 6)
Independent Evaluation of the Impact of the Assessment System

Independent evaluations of the effects of any national assessment system must be conducted on a continuing basis. These evaluations should be conducted by one or more organizations independent of any group chosen to establish National Standards, or any group which develops or approves assessments used in the national assessment system.

Technical Assistance, Development, and Research

To assess the National Education Goals, we need “Break the Mold” assessments. Substantial federal financial support must be available to states and local districts, working in partnership with commercial test publishers, universities and other educational institutions, to develop new assessments that fulfill the vision of a positive assessment role in educational reform.

The national government must provide technical assistance to states, local school districts, and their teachers to help them profit from the planned value of the assessment system.

Finally, a research program should be funded to solve problems in design, analysis, and interpretation of new assessments. Areas requiring research should include, among others, new approaches to comparability, generalizability and fairness of assessments, adaptation of assessments to special needs of students, and issues of aggregation and reporting.

All of these functions could be carried out by the U.S. Department of Education.

Issues in the Implementation of the Proposed System

Incentives. What incentives for participation in assessment development and implementation must be created?

An obvious choice is to relieve schools from certain federal and state testing requirements. Logically, we cannot imagine simply layering more assessment onto that which is already in place. We suggest that decisions about incentives should be an early agenda item for deliberation by any successor group to the Council.

Costs. The proposed system costs money and reallocation of teachers’ and administrators’ time. By any rough estimate the system will be expensive. Costs will depend upon frequency of assessment tasks, whether scoring is internal or external to the school. Sampling students could vastly reduce costs.

Who Pays for It? State and local governments are not able to support new assessment costs at the same time state economies are in trouble. Support is essential by the federal government for the development of “Break the Mold” assessments by states or groups of states to meet the National Standards.

We may not be able to pay the full bill right now, and we may want to pay as we go. States may very well have to size their efforts to the pocketbook available and pick particular subject matters on which to concentrate.

Implementation Strategy. We propose taking our lead from the business community and adopting an implementation strategy that rests on the idea of rapid prototyping—getting something, even something modest, into trials early, and adopting a test-and-fix approach—to improve our assessments as we go. Critical to this overall strategy are three points: coordinating development in a few priority areas, for example, mathematics, reading, and writing for fourth grade; encouraging different strategies and approaches; and reporting descriptive information as early as possible to the public, parents, policymakers, educators, and students.

First Steps Recommended to the Council

- Encourage the development and trial of assessments of the NCTM standards.
- Make governance decisions and seek authorization and funding for the National Quality Assurance function.
- Encourage the funding of research projects and new “Break the Mold” assessments.
- Move toward encouraging the relief of certain existing federal and state testing requirements for participants in the National Standards process.

Assessment Task Force

Eva L. Baker, Chair
National Center for Research on Evaluation, Standards, and Student Testing, University of California, Los Angeles

David Kearns
Deputy Secretary
U.S. Department of Education

Edward Meyen
Dean, School of Education
University of Kansas

Joan Baron
Common Core of Learning Assessment Program, Connecticut Department of Education

Curtis Banks
Professor of Psychology
Howard University

Dale Carlson
California Assessment Program
California Department of Education

Georgia Earnest Garcia
Assistant Professor
University of Illinois

Edmund Gordon
Professor Emeritus, Yale University

Daniel Koretz
The RAND Corporation

(continued on page 7)
the following technical reports have recently been released and are available through the CSE/CRESST office. To order any report or for a complete listing of all CSE/CRESST reports, monographs and resource papers, please contact Kim Hurst, (310) 206-1532.

**Writing Portfolios at the Elementary Level:**
* A Study of Methods for Writing Assessment  
  Gearhart, Herman, Baker & Whittaker  
  CSE Technical Report 337, 1992 ($4.00)

This study examines the feasibility of using students’ writing portfolios to evaluate writing competence. Analytic portfolio ratings showed promising levels of measurement quality, but differences in assessed level of performance emerged when portfolio scores were compared to other assessments. Qualitative analyses of the scoring process revealed significant design challenges, particularly in devising portfolios which reflect classroom instruction yet are sufficiently uniform to permit meaningful comparisons within and between classrooms and schools.

**A New Mirror for the Classroom: A Technology-Based Tool for Documenting the Impact of Technology on Instruction**  
  Gearhart, Herman, Baker, Novak & Whittaker  
  CSE Technical Report 336, 1991 ($5.00)

Since 1987, UCLA’s Center for Technology Assessment has been conducting a set of evaluation, research, and development activities at selected Apple Classrooms of Tomorrow (ACOT) sites, with the goal of documenting the impact of technology access on K-12 environments. This report includes a description of a new measure, a technology-based classroom observation instrument for documenting the impact of technology on classroom instruction.

**Cross-State Comparability of Judgments of Student Writing:**
* Results from the New Standards Project Workshop  
  Linn, Kiplinger, Chapman & LeMahieu  
  CSE Technical Report 335, 1991 ($5.50)

This study explores the feasibility of comparing student performance on different writing tasks across states, a critical question in current debates about a national examination system. The report is based on data from ten states that participated in a cross-state scoring workshop. Results show a high level of agreement as to what constitutes low-to-high quality writing.

According to this report, however, “considerable cross-state discussion will be required to arrive at common performance standards.”
Retirements, Awards, Memorials

Wanetta Jones Retires

We are sorry to say good-bye to Wanetta Jones who worked at the Center for the Study of Evaluation since 1979 and at CRESST since 1985. Wanetta’s responsibilities included managing the Visiting Scholars Program, consultant agreements, and various national, state, and local conferences. Just barely retired, Wanetta said that she is looking forward to an active travel schedule as a consultant. We wish her the best in her new endeavors.

Darrell Bock Receives Award

Congratulations to Darrell Bock, University of Chicago, who recently received an award for “Distinguished Contributions to Measurement.” Professor Bock is a CRESST partner (National Opinion Research Center) and longtime friend. The award was based on the votes of the presidents of: the Psychometric Society, Division D of the American Educational Research Association, Divisions 5 and 15 of the American Psychological Association, the Association for Measurement and Evaluation in Counseling and Development, and the National Council on Measurement in Education. Professor Bock received the award at the Education Testing Service Invitational Conference in October, 1991.

T. Anne Cleary Memorial

It was with tremendous sadness that we learned of the tragic death of our friend and colleague, T. Anne Cleary, Associate Vice President of Academic Affairs, University of Iowa. The T. Anne Cleary Psychological Research Scholarship has been established in her honor. Monetary awards will be available each year for one international student and one native born student in the advanced stages of doctoral work in the field of psychological and quantitative foundations. Donations may be sent to: The T. Anne Cleary Memorial Scholarship Fund, University of Iowa Foundation, Iowa City, IA, 52242. Or call the foundation at: (319) 335-3305.

Additionally, the American Psychological Association will hold a special symposium called “Maximizing the Meaningfulness of Measuring: A Tribute to T. Anne Cleary.” Featured speakers are Robert Linn, CRESST/University of Colorado, Richard Snow, Stanford University, Carol Tittle, City University of New York, Nancy Cole and William Angloff, Education Testing Service. The annual APA meeting is August 14-18, 1992 in Washington, DC.