Questioning Assessment:
How Do We Learn What They Are Learning?

by Kathleen Cushman

The notion of “performance assessment” comes up again and again, as Michigan educational policymakers wrestle with new definitions of student success and ways to measure it. Schools must begin, the theory goes, to ask students to demonstrate in concrete and “authentic” ways—not just through paper-and-pencil, conventional tests—what they know and are able to do. Then teachers can use that information not only to assess student understanding and mastery, but to restructure their curriculum and teaching practices toward improving their knowledge and abilities.

Whether such an effort starts out at the top or the bottom of the educational bureaucracy, whether it defines subject-area “curriculum frameworks” or crosses the disciplines to describe “learning outcomes” in terms of thinking skills, in the end it is bound to end up in the same place—the classroom. And increasingly, educators are realizing that unless classroom teachers are intimately involved in designing their own performance assessments, change will be slow and uncertain at best.

This issue of Changing Minds explores ways that different schools and teachers—both from Michigan and from around the country—have taken that challenge into their own hands. Rather than depending on off-the-shelf tests, these teachers have looked hard at their own goals—what knowledge and abilities they want their own students to have when they leave this course, this school—and revised their assessments to emphasize open-ended questions, problem-solving strategies, essays, and other “performance” tasks.

Inevitably, the old attitude that a test is something to take after you learn something begins to give way to a new approach: the belief that assessment is an integral part of day-to-day learning, often almost indistinguishable from regular class time. And teachers in these classrooms have begun to design their students’ daily lessons to allow practice in the higher order skills that old-style tests ignore—closely reading and discussing sources, coming up with and researching ideas, solving complex problems, initiating projects, and creating original work.

A Collage of Examples

In San Diego we will visit a middle school where the entire student body takes on the same quarterly performance task—this time, an analysis of the relation between philosophies of nonviolence and last year’s uprising in South Central Los Angeles. This school’s teachers are closely involved with state and national efforts to revise assessment practice; but rather than depending on outsiders to tell them what to do and how, they are drawing the best ideas they can get from those outside resources and applying them in the ways that fit their own situation. In this case, they want a whole-school snapshot of how far students have traveled on the road to written and mathematical literacy—and
they want to take the scored tests right back to the students for revision.

We will look at a school from suburban New York (historically one of the most rigidly over-tested states in the country) where, encouraged by a new state education department compact, teachers are devising new ways for students to earn the prestigious Regents diploma. With support from both the district and the state — in the form of time, money, and professional development opportunities — more and more teachers are trying out new assessment methods, fine-tuning them with the help of "critical friends," and beginning to build new standards that truly arise from their community's expectations.

In Vermont, we will look in on a group of fourth- and eighth-grade mathematics teachers who are meeting to hash out the details of an ambitious new statewide portfolio assessment system. How teachers can first identify good portfolio problems and then score them reliably and consistently has been a major challenge. Vermont has responded by establishing teacher-run networks that aim to involve every teacher in the state. Visiting committees validate each school's scores by rescoring random portfolios; and the state supplements this information with a conventional test with which to compare districts' progress.

In Michigan's Holt High School, five years of working as a professional development school have refocused the faculty's attention on whether students are experiencing conceptual changes — that is, really understanding how knowledge is applied, rather than merely being able to regurgitate information they acquire by rote. In classes throughout the school, teachers have designed their assessments to answer that question, using everything from personal interviews to outside panels in the effort.

**Teachers Keep Learning, Too**

In the current trend toward state and national assessment redesign, many policymakers ignore a crucial task — to provide a continuous learning opportunity for teachers, not just for students, so as to generate local energy and invention in the interests of all students' achievement. In fact, whatever the "test" looks like, it works less well when it arrives from an outside source. Only by working from the inside out can teachers embed ongoing assessment tasks in the curriculum itself, rather than treating them as separate from the rest of teaching and learning. As in real life, where we assess ourselves continually in striving to achieve tasks we care about, the assessment is not an isolated occurrence that sorts and selects our winners and losers, but rather an honest, authentic feedback device coaching us along the path to excellence.

This issue's collage of examples from around the country is meant to inspire teachers to start reflecting on how they too can create more revealing and useful classroom assessments. We interview teachers who develop assessments with each other and with their students. For just as teachers too often get left out of assessment design, so do students. We provide examples of how learning to evaluate their own work can ultimately give students a strong internal sense of what high standards mean.

**When faculties begin to scrutinize their students' work and discuss whether it meets their own standards, they have taken the first step toward changing curriculum.**

When faculties begin to scrutinize their students' work and discuss whether it meets their own standards, they have taken the first step toward changing curriculum and pedagogy in meaningful ways, and recasting "accountability" as a positive thing, not a big stick from the outside. Because they know their students well, teachers have far greater potential than does any standardized test to assess them in rich and meaningful ways. Community by community, we see them here working to achieve high standards without accepting standardization.

Editor: Kathleen Cushman

The guest editor of this issue of Changing Minds, writes HORACE, the journal of the Coalition of Essential Schools, a national school-reform movement based at Brown University in Providence, Rhode Island. Her work has also appeared in many national magazines.
"In This Case, Teaching To The Test Is Good"
Another Way to Evaluate Schoolwide Progress

You would never guess, visiting on a sunny day in January at the O'Farrell Community School in San Diego, California, that the 1,400 middle school students here are in the middle of an important end-of-quarter schoolwide test.

To the unsuspecting eye, it is an ordinary day in a lively classroom. In Lenora Smith's humanities class, 30 seventh and eighth graders sit around tables and talk about the Los Angeles uprising that followed the 1992 Rodney King verdicts. Their aim after a week of background readings, videotapes, and brainstorming: to connect their own impressions of the L.A. events to the philosophies of nonviolent resistance preached by Dr. Martin Luther King, Jr. and India's Mahatma Gandhi. The next day, each student will write a letter to another student in South Central Los Angeles, conveying their own understanding of the tension between social injustice and the principles of nonviolence.

It looks less like a test than like an engaging and demanding unit in this school's theme-based curriculum, which centers schoolwide on a quarterly, cross-disciplinary "essential question" — the device educator Theodore Sizer suggests, in his groundbreaking book Horace's Compromise, as a way to deepen and enrich student inquiry skills. (This quarter's question at O'Farrell: "Where does conflict lead?".) It will take over a week. No secrecy is involved; the task is shared and discussed, and every student knows the criteria for doing well. There will even be time for revision after the task is scored. Not what you would expect from yesterday's schools, perhaps—but this is O'Farrell's way of measuring its students' progress toward written self-expression—a serious, rigorous assessment that will immediately inform all teachers' classroom practice.

Knowing How All Students Are Doing

O'Farrell teachers obtained this particular assessment because a group of its teachers works closely with the New Standards Project, a heavily funded research effort involving 18 states and headed by Lauren Resnick at the University of Pittsburgh and Marc Tucker at the National Center on Education and the

---

**Math Concepts**

**Self-Evaluation**

Please attach this sheet to all of your quizzes. Also attach any notes or assignments that could have helped you on this quiz.

1. How many questions did you answer correctly?

2. How many questions were on the quiz?

3. Divide the numbers above to get your score?

4. Did you feel prepared for the quiz?

5. How could you have prepared yourself to do better on this quiz?

6. Write in the problems that you answered incorrectly:

   A. 
   B. 
   C. 
   D. 
   E. 
   F. 
   G. 
   H. 

7. Briefly describe what mistakes you made in each of the problems above:

   A. 
   B. 
   C. 
   D. 
   E. 
   F. 
   G. 
   H. 

8. How will you go about learning how to do these problems correctly?

   __________________________________________

WHEN APPROPRIATE YOU MAY FASTEN THIS TO OTHER PROJECTS OR ASSIGNMENTS WHERE YOU HAD TO USE THESE SKILLS.
Economy. New Standards hopes to fashion a national examination system based on performance assessment, which would allow states and districts to assess programs based on how their students do in comparison to others on specific tasks. Usually only a sampling of students from a school would take a New Standards "test"; the tests would be scored by outsiders; and the results would eventually make their way back to the school.

"The glory of our efforts will be evidenced when we take what we've learned back to our students and creatively and collaboratively encourage them to produce qualitatively better finished products."

"We wanted a truly seamless web of teaching, learning, and assessment," says teacher Byron King. "Schools need to use nationally developed and calibrated assessments to serve their own classroom needs—to shape what they need and want. It's a whole different kind of relationship between the national testing people and the school."

**Scoring Across Disciplines**

This whole process of school-wide assessment is repeated every quarter at O'Farrell; students in every grade attack the same two performance tasks in mathematical reasoning and written self-expression. Each task has a clearly defined, measurable goal. The project about nonviolence, for instance, charges students to "demonstrate their writing ability to convince or influence others in a persuasive essay." The math task for this quarter—which teachers devised themselves—took much less time than the writing task, but it also reflected the schoolwide theme of conflict: To show they could "analyze the possible outcomes of an event," students had to reason mathematically whether the rules of a coin-throwing

---

**Portfolio Process: Self-Evaluation Step**

**Personal Reflection**

1. Which challenge(s) of The O'Farrell Standard does this piece address?

2. What was (were) the expectation(s) for this assignment?

3. Do I feel that I have met the expectation(s)? (Mark an "x" on the line below)
   - Below
   - Meets
   - Exceeds

4. How do my results compare to the expectation(s) of the assignment?

5. What do I feel I learned through creating this piece?

6. What strengths do I feel are exhibited in this piece of work?

7. If I had more time to work on this piece, I would... because...

8. I would like feedback on the following... (Check back: List 3 areas or 3 questions to be answered!)

**Comments from Guests & Visitors:** Please write down any comments/suggestions/compliments if you should happen to look at some of my work. Please write the date and sign your name. Thank you.

---

Figure 2
Devising their own tasks makes O'Farrell teachers keenly aware of what their expectations are for students.

game were fair, then suggest ways to alter them.

Why and how should math and writing tasks be scored by teachers not trained in that particular field? "We want every teacher to become more aware of how important mathematical reasoning and written expression are across the curriculum," says Lenora Smith. But she concedes that it takes time and training for teachers to learn fair scoring techniques. "If you're not familiar with children's writing, you might have trouble getting past mistakes in the conventions of grammar and spelling to be able to judge the content mastery and rhetorical skills the student displays." From California's Learning Assessment System (CLAS), she says, "We stole a two-part rubric that scores conventions and rhetoric separately."

When teachers gather to score the math task, they work in small groups using a clear description of what must be communicated by a successful answer to the coin-toss problem. Then they score each with a number from 1 to 3, based on whether the student (1) needs more instruction either individually or in small groups, (2) is ready to revise with a simple comment or clarifying question from the teacher, or (3) far exceeds expectations. Teachers confer with colleagues when they are perplexed about what grade to give, but generally the range of "twos" is very broad, says O'Farrell team leader Becky Breddlove, who devised the rubric.

"We consider all papers to be rough draft first attempts that can be improved upon," she says. "The glory
O'Farrell teachers who work closely with another national assessment project, the Performance Assessment Collaboratives in Education (PACE), based at Harvard University. Headed by Dennie Palmer Wolf, PACE emphasizes portfolio assessment as a way to follow students’ “footprints” toward eventual mastery. By including in their portfolios work at different stages — and by reflecting in writing about where it has led them — students begin to take ownership of their own learning process as well as the final product.

Yoshida, for example, now supplies a self-evaluation sheet that students may attach to every quiz, project, or assignment they turn in to him (see Figure 1). Meant to help them analyze where they went wrong and what they can do to prepare to do better, the practice specifically asks students to take responsibility for their own learning. Schoolwide, students have comparable involvement in their own assessment: Every portfolio piece and quarterly assessment has a “personal reflection” section, in which students identify what challenges and expectations it posed, evaluate how they met these, and spell out what else they might need to go further with the task (see Figure 2).

Portfolios reduce the school’s reliance on any one performance assessment task as a final measure of student success, Byron King points out. And California encourages schools to use them as a self-assessment as well; the state trains faculty members in a systematic protocol through which they can examine student work and determine what steps the whole school needs to take toward higher achievement.

With all this rich array of assessments, how does a school manage to boil down student progress into an efficient report card format? At O'Farrell, well supplied with computer technology through grants from Apple, Panasonic, and other corporations, the answer lies on disk. All teachers submit quarterly progress reports describing student progress toward specific elements of the “O’Farrell Standard,” a list of performance requirements that range from community service to research projects, from demonstrated academic competence to behavior and attitudes. The system of educational “families” at this school, in which a “home-base” teacher serves as primary advisor for around 25 students, aims to ensure that teachers look closely at every student’s overall progress and comment on it in writing.

Two teams have taken the computerized reports to another level altogether, creating a digital portfolio in which students themselves place and evaluate their own work. Using Hypercard stacks to relate each entry to the O’Farrell Standard, the students devise a vivid electronic picture of their achievements, which can include everything from a wrap-up of their research thesis to a biographical poem, a reflection on the schoolwide “essential question,” and a thoughtful analysis of their intellectual strengths, weaknesses, and ambitions (see Figure 3).

“Their folder may belong to the teacher,” says technology teacher Don Beck, surveying the array that resulted from his second year of getting students to create their own digital portfolios. “But this belongs to the kids. I'm completely blown away by their creativity.”

For more information contact:

O'Farrell Community School
Center for Advanced
Academic Studies
6130 Skyline Drive
San Diego, CA 92114
(619) 263-3009
Taking Back Testing: A School Begins to Shape New Standards

There was a time not long ago, at Croton-Harmon High School in this rural Hudson Valley town an hour north of New York City, when the dreaded Regents examinations arrived in locked boxes on the appointed year-end day, to test the wits of those students deemed advanced enough to try for the state's prestigious Regents diploma. New York had earned a national reputation for the frequency and rigor of its testing, and these course-end exams had the same kind of stranglehold on the curriculum as, for example, an Advanced Placement test out of Princeton.

But something has begun to change in New York in the last year, and in a dozen different ways you can see it in the classrooms of Croton-Harmon these days. Teachers long accustomed to outside control of what they teach and how they test are devising their own assessments. And instead of bucking the state's regulations to do it, they have the backing of everyone from the superintendent to the Board of Regents itself.

In fact, under the state's New Compact for Learning, priority is now going to giving local schools and districts the authority to determine what they want their graduates to know and to be able to do—and what evidence they will accept as proof. Commissioner of Education Thomas Sobol has begun to revise the old Regents exam system dramatically, substituting a new philosophy of state support for local assessment choices and focusing on individualized performance assessments such as portfolios and exhibitions. If Sobol has his way, no longer will the state encourage separating out some students to fulfill the higher expectations of classes leading to a Regents diploma; instead, all students would complete a Regents portfolio that demands more of everyone.

The decision shakes the very foundations of a district like Croton-Harmon, a middle-class, mostly white exurban community with an uneasy hunch that its high-achieving students might nonetheless be graduating unprepared to compete in a 21st century global economy. Inspired by Commissioner Sobol's vision, several years ago its high school principal, Sherry King (now the district superintendent), went to community focus groups to learn more about the public's ultimate learning goals for students.

"We found surprising agreement," King says. "Parents and community members all wanted to see students demonstrate mastery in more concrete ways than standardized paper-and-pencil tests can possibly provide." Drawing on that input, as well as on advice from local experts and university people, on the standard-setting work of national professional groups, and on the assistance of the Coalition of Essential Schools, the Croton-Harmon teachers began to hammer out their own ways of measuring student progress in new and more meaningful ways.

"I decided to incorporate a substantial oral element into my weekly quizzes." A 'Dinosaur' Diverges

"I'm a dinosaur, and a damn good dinosaur," boasts high school chemistry teacher Ron Lavery, perched laughing behind the tall lab table covered with student records. After 33 years preparing Croton-Harmon students for their chemistry Regents exam every year, Lavery was already well seasoned in older assessment techniques four years ago when principal King's challenge to try different approaches caught his imagination.

"First, I decided to incorporate a substantial oral element into my weekly quizzes," he says. "Besides improving their oral skills, I also wanted the kids to take more responsibility for their own learning. Instead of their always coming to me, I wanted them to be able to explain concepts in chemistry to each other." To cover part of every chapter, Lavery now hands out a list of a dozen questions to the class and asks students to prepare the answers with a partner. ("If they have problems, they can only come to me after consulting first with their partner and then with someone else in the class," he says.) The quiz consists of each team's oral answers to two of the questions, which Lavery picks from the list at random; the two students may refer to their notes (but not simply read from them) and will receive the same grade. "I use a double period for this about once a week, on average," he says. "I found it a very effective way of covering relatively uncomplicated material."

The success of his experiment led Lavery to a more ambitious idea: replacing his course's midterm exam with an oral equivalent. Now he blocks off three days in January to conduct 15-minute sessions with each of his 80 students, in which he questions them closely to draw out their understanding of the subject matter. (Students must prepare to an-
swer discussion questions across the board; they then draw two index cards with two questions on each from the array on the teacher’s desk."

“It’s a very inefficient use of my time,” Lavery concedes of the marathon schedule that results. “But I believe in it a hundred percent, no matter how tiring it proves.” The one-on-one element, he says, conveys information about student progress that written testing simply cannot achieve. “You can’t cheat, you can’t run and you can’t hide,” he says. “At the same time, there’s a friendly, Socratic feeling to the process. There’s not necessarily one right answer; I use follow-up questions, pursuing a topic until I get a strong sense of how well the student gets it. They can pause, sit, and think. A lot of kids listen, read, and comprehend the material fairly well, but aren’t always able to translate their understanding onto paper. I give them the benefit of that doubt, designing my questions to extract that understanding if it’s there.”

Like a gymnastics judge at a performance, Lavery takes notes but does not rely on a written rubric to assign his score. The oral exam makes up 50% of the midterm grade—the rest derives from homework and quizzes—but this teacher sets high store by it nonetheless. In general, his records over four years show, students score better on their oral than on their written assessments. And further corroboration comes from written student reflections on their own performance before they get their grade.

“I’ve had students tell me they were surprised they knew that much,” he says. “It gives me a wonderful, powerful feeling—I knew this kid knew it! He did understand, and this test discovered it.” The next step, he believes, is to bring students back a year or two later and see how much they remember. “The wonderful thing about chemistry is that it answers a lot of ordinary questions that come up in everyday life,” he says. “Can you tell me, for example, why they have special instructions for high altitudes on the sides of cake mixes?”

This year Lavery will have the chance to adjust the state's traditional Regents exam to reflect his new assessment methods. Though his students will still take Part 1 of the exam for 65% of their grade, the state will allow him to substitute a third-quarter research paper for 15%, a lab practical exam for 10%, and an oral component Lavery designed for the final 10%. It is just one example of how, in a dramatic shift of control, a tightly regulated state exam system is giving way to a new ethic of trust and empowerment for local schools and communities. This veteran teacher—a self-proclaimed “dinosaur” because of his insistence on content coverage—has developed a contagious energy and commitment toward new ways of measuring student understanding. “I don’t see why this method shouldn’t work in any area,” Ron Lavery declares.

Surviving Doomsday

Upstairs in the 9th- and 10th-grade science classroom, teachers Judy Kelley and Jean Kennedy mull over their own alternatives to the Regents earth science and biology exams. “Traditionally the Regents had 14 topic units, which many teachers found impossible to prepare for adequately,” says Kelley. “Some years ago it was revamped to have fewer topics and more choice, and now the state lets you apply to replace 25 to 30% of the test with an alternative assessment of your own design.”

With this in mind, Kelley devised the Doomsday Dilemma, a journal project that spans three of the year’s quarters yet counts for 20 points on the final Regents exam. “Assessment shouldn’t be something you do at the end of the course,” Kelley says. “It should be used throughout as a tool— to coach students toward the same high standard, to give them the chance to correct themselves, and to link the disciplines by rigorous habits of mind.”

Every student in Kelley’s four classes starts the year in a race against time: Because of a catastrophic comet hurtling toward Earth, they must decide which of four fictional alternatives “Regenesis” planets might support human civilization. Working in teams assigned to each planet, they evaluate rock samples, temperature data, shadow angles, and seismic data obtained from exploratory probes, and keep a scientific journal as their deductions unfold. They work through the modest data given, sketching and analyzing their observations, making inferences and predictions, and defending or revising their conclusions in response to Kelley’s written comments.

“At the end of the third quarter,” Kelley says, “the teams from all four classes get together and present to the world a picture of their planet’s good points, weaknesses, and ways around those weaknesses.” The audience includes not only their peers but interested teachers, parents, and community members invited to witness and question their conclusions. “They will discover that each planet has its fatal flaw,” she says. “Perhaps its density is twice that of Earth’s, which makes for an intolerable gravitational force. Or perhaps the tilt of the planet in relation to the sun is not enough to sustain survival.”

Kelley gave extra credit this year to students who could orally explain to her the tilt of their imaginary planet; and then she went one step further. “They could get even more points if they could explain it to another student who could then satisfactorily explain it to me,” she says with barely suppressed glee. “It was amazing how that strategy leveled the
playing field in a class that isn't grouped by ability."

In addition, she says, this project's scientific richness allowed these earth science students to interweave the course's various content areas as they tried to solve their problem. The journal allows continuous assessment of their mastery in several learning areas at once—content, scientific skills, and communication and teamwork.

Tailoring the Regents for Biology

For her colleague Jean Kennedy, who has taught Regents biology at this school for 35 years, Kelley's one-project approach did not work. "In one year I have to cover evolution, human physiology, plant and animal structure and function, reproduction and genetics, and ecology," she says. "I just couldn't find a project rich enough to support sufficient student learning in all those areas."

Instead, she decided to tailor the Regents exam so that students could replace a multiple-choice section counting for 20% with a written research report, an exhibition, a community service activity, or a portfolio that demonstrated mastery of one extended area of the curriculum. Students' plants line the walls of her classroom, and these young scientists chart and observe them as tenderly as infants; after they framed their hypotheses, completed their investigations, and had their final presentations assessed, almost half the class chose to continue their observations for the rest of the year.

In addition, Kennedy replaced the practical section of the Regents, containing short-answer questions on lab skills, with performance tests in the lab throughout the year. Wherever she observes student mastery developing, Kennedy keeps an ongoing evaluation record that will count for 15% of the final Regents grade.

Finally, this teacher puts her money where her mouth is when it comes to incorporating assessment into daily classroom activities. "I noticed that students often throw away their homework or lab reports after they've been checked," she says. "In order to show them the importance of this preliminary work, 1 allow them to use their own past work when they take a test." This creates a more authentic situation, Kennedy believes—more like the experience of real-life scientists—and lends purpose and weight to otherwise routine assignments.

A Meeting of Minds

Shelly Frisch and Julie Henderson teach 10th graders in the second year of a humanities course called Global Studies, for which the Regents board also schedules a culminating exam. "But I thought the exam trivialized the subject matter," says Frisch. "We got the idea to replace it completely, with an exhibition that would pull together the history and literature of several eras and areas."

Last year for the first time, the Global Studies class set aside five days for "A Meeting of the Minds," a public presentation held in the school auditorium in front of classmates, parents, community members, and other outside visitors. During each 40-minute presentation three students who had taken on the identities of historical figures ranging from Joan of Arc to Nelson Mandela, Charlemagne to Vaclav Havel were on stage. The discussions that followed were lively debates about a social, political, or economic issue—such as "Is the situation in Bosnia a 'just war' ?" or "Does the government have the right to regulate individual morality ?"—from their different points of view. The unit's essential question overarches the entire investigation: "Have the times made the individual, or has the individual made the times?"

"Five weeks beforehand we stop all other work to prepare for the exhibition," says Frisch. "Every student has to become quite familiar with each of the three characters her team chooses, and the three must both span geographic boundaries and cross historical epochs. To do justice to our ninth-grade curriculum, we also require that one character be from a developing country." After showing sufficient understanding of all three by passing a "fluff check" on factual matters, the students do more thorough research into a single personage, preparing a written bibliography.

"It's amazing how absorbing and important the project becomes to them," says Henderson. "Students have already started to adjust their personal plans so as not to interfere." On the big day, the 30-minute presentation is followed by 10 minutes of questioning from the audience; finally, students must write a critical essay analyzing the course's essential question using evidence gathered from watching a classmate's exhibition. Henderson and Frisch score each exhibition using a detailed rubric that weighs work habits, research and development, presentation, speaking skills, and the bibliography as well as the written essay's content and style.

Making Accountability Local

What strikes the visitor in talking to Croton-Harmon teachers is the diversity and flexibility of their approaches to alternative assessments. This district—and this state—has clearly given practitioners the message that variety and authenticity mean more than a uniform testing code. These teachers are working together to make connections between disciplines; they hash over the habits of mind they want to instill in every graduate and in every subject area; they often critique and serve as checks on each other's ideas and practice. And where this requires a change in the school structure—altering the schedule for longer blocks, for instance, or allowing time for
An Ancient Puzzle and A Modern Challenge
Vermont Teachers Craft Their Own Portfolio System

Outside the windows of this country inn high in the mountains of Randolph, Vermont, the autumn hills blaze with color against a brilliant blue sky. In the meeting rooms inside, the intensity is no less striking. The thirteen eighth-grade math teachers who have gathered here from around the state have an ancient puzzle to solve among themselves, and a thoroughly modern charge to carry out.

The puzzle, known as the Tower of Hanoi, was brought in by one of these teachers as an example of how she engages students across a wide achievement range. On the big conference table she has put it out: a set of 20 disks graduated in size, which are arranged with the biggest one at the bottom on the first of three spindles, not unlike a Playskool donut toy.

Just as the ancient monks of Tibet did some 2,000 years ago, these teachers are trying to discover the minimum number of moves to relocate the disks, one at a time, from spindle A to spindle C — without ever placing a larger disk atop a smaller one, and with the aim of reproducing the original pile on spindle C. (see Figure 4). The monks thought that the world would come to an end if they could solve the problem for 64 disks; these teachers are slightly less portentous, but equally serious about the importance of their task.

"I just left it on the back shelf of my classroom for weeks," says Rita LaPier of Brown River Middle School in Underhill, Vermont. "It's impossible to let alone—kids kept going back to fool around with it." Huddled in fascinated groups of two and three, these teachers quickly start trying different approaches; before long they have discovered that it takes a minimum of 7 moves to move even three disks across, and some are well on the way to extrapolating a mathematical rule that will govern the answer for 20.

Ancient or modern, it is an ideal problem, they decide as soon as they can tear themselves away. Charged by their state with coming up with examples of the kinds of mathematical challenge that ought best be included in Vermont's ambitious new portfolio assessment system, these teachers agree that the Tower of Hanoi meets most of their criteria.

One by one they peruse the criteria posed by the National Council of Teachers of Mathematics (NCTM) in defining a "worthwhile mathematical task." The Hanoi Tower is certainly "sound and significant mathematics," they agree; the puzzle element draws on students' understandings and interests; and it clearly engages the intellects of learners at different levels and in different ways. It develops understanding of mathematical ideas, and it could stimulate students to make other connections — from Morse code to chain letters — that follow similar binary patterns. In trying to solve it, students could break the problem into smaller parts or apply it to bigger numbers. And finally, it prods students to communicate with each other; in explaining how they got their results, students learn that using variables, functions, and other scientific notation can

For more information contact:
Croton-Harmon School
Old Post Road, South
Croton-on-Hudson, NY 10520
(914) 271-2140
simplify their sentences. All that makes the Tower of Hanoi a "good portfolio problem," they decide — an authentic, "real" problem that would render a Vermont student portfolio more meaningful than any multiple-choice conventional measure. Drawing on Marilyn Burns's book, About Teaching Mathematics (Sausalito, California: Math Solutions Publications, 1992), they work through the criteria for "real problems":

1. A perplexing situation that the student understands is presented.
2. The student is interested in finding a solution.
3. The student is unable to proceed directly toward a solution.
4. The solution requires use of mathematical ideas.

These criteria, they decide, certainly vary from student to student, problem to problem. But the Tower has potential. They turn to some student solutions to chew over the next challenge: How will the work be scored?

Right Answers Aren't Everything

"Portfolios help kids understand that we're interested in the process of learning, not so much the right answer every time," says Clare Forseth, a sixth-grade math and science teacher at Maritan Cross School in Norwich, Vermont and a member of the original team that developed the new system. "We want to change the way math is being taught towards more problem solving and communication." To assess this, teachers devised seven criteria, the spine of the rubric teachers and students both would learn to use on a regular basis.

To assess problem solving, for example, they ask:
- Whether students understood the task
- How they solved the problem
- Why they made various decisions along the way
- What extensions they could make from their solution – observations, connections, or general rules.

To assess communication, they look at:
- The use of mathematical language (vocabulary, symbols, and notation)
- The ability to use mathematical representation (graphs, charts, tables, diagrams, and models)
- The clarity and organization of the presentation.

Four levels of competence are spelled out in each of these areas in the scoring rubric Vermont teachers have devised; they also have selected "benchmark" samples of student work that exemplify each level of performance in each of the criteria. Students too must learn to evaluate their own work by the same criteria. (see Figure 5). Still, teaching teachers to score student work in ways that consistently agree has been a major challenge of this ambitious statewide effort. The teachers in this group, for example, have just finished a lively debate about how to score student work that is clearly "scribed" by an instructional assistant for a student with special needs. (The education department's response: Make no distinctions about such pieces, labeling someone's work as "special ed" encourages holding different children to different standards.)

Back at the Tower of Hanoi, the Randolph group is ready to look at a few students' solutions with the rubric's criteria in mind. "Most students start by making up a table of the number of moves and then

The Tower of Hanoi

A legend about a group of ancient priests somewhere in the world tells of them solving a very old puzzle called the Tower of Hanoi.

The Tower of Hanoi is a set of three spindles or poles upon which are placed disks of descending sizes. Three such disks are shown in the picture below.

The puzzle or challenge is to move the tower of disks to a different spindle following two rules:
1. You may move only one disk at a time.
2. At no time can a larger disk be placed on top of a smaller disk.

Using these rules, see if you can figure out how to solve the three-disk puzzle. You may want to get three different coins to represent the three disks and lay out three spaces on a piece of paper.

Figure 4 - Reprinted by permission of the publisher
noticing a pattern that can be translated into a mathematical sentence," says the teacher on my left. "Look at this one." We all stare at the piece of student work: three columns (number of objects, number of steps, difference between the number of steps necessary to move three objects successfully and the number necessary to move four objects. The student has written out a fairly lengthy explanation of his procedure ("I looked for a pattern or something that could help in creating a formula"), and sure enough, he has come up with a mathematical sentence he thinks can describe the progression (2n + 1, where n equals the number of steps).

For five or six minutes the teachers fall silent, scoring the student’s performance according to the 7 criteria described. Finally they look up. Going around the table, they compare their scores in each category — displaying a high degree of agreement, but not unanimity.

"He's got the wrong answer," someone protests. I have been wondering about the same thing. I look at the rubric in front of me: no category for "got the right answer." Am I watching the Emperor’s New Clothes of the new math world?

“That’s true,” says teacher Mark Johnson, who teaches at the Main Street Middle School in Montpelier. “But look at all the really important things this kid did master. He’s using problem-solving skills — setting up a table, explaining his decisions clearly, moving on to a general rule. His communication is good — well organized and detailed, with a clear table, and math notation in the general rule he devises. And he’s exploring important areas of content that we’re concerned that he know: patterns and functions, powers, the binary system. So what if he doesn’t get the right answer the first time? We know he has all the tools to try again.”

For all these reasons, the group agrees, the Tower of Hanoi makes a good portfolio assessment problem. Talking their scattered differences through, they even agree about how to score this work. But that isn’t the end of their challenge, they observe. What kinds of problems should a class use to lead up to being able to tackle the Hanoi Tower? How should they use it to begin to explore new areas? Authentic assessment should be folded into the curriculum, not set apart from it, they agree; it should provoke new learning, not just wrap it up.

**Vermont Mathematics Criteria Student Self Assessment**

**PS 1: Understanding the Task**

1. I didn’t understand enough to get started or even make progress.
2. I understood enough to solve part of the problem or to get part of a solution.
3. I understood the problem.
4. I identified special factors that differentiated the way I approached the problem.

**PS 2: How you Solved the Problem**

1. My approach didn’t work.
2. My approach would only let me solve part of the problem.
3. My approach would work for the problem.
4. My approach was efficient and/or sophisticated.

**PS 3: Why- Decisions Along the Way**

1. I had no reasons for the decisions I made.
2. I knew I was reasoning but it was hard to see clearly why the reasoning was used.
3. Although I didn’t explain the reasoning for the decisions, my work suggests the reasoning was used and made along the way.

**PS 4: So What - Outcomes of Activities**

1. I solved the problem and then stopped.
2. I solved the problem and then made comments about something I observed in my solution.
3. I solved the problem and connected my solution to other math that I know OR I described a use for what I learned in the "real world."
4. After I solved the problem I thought about the solution.

**C 1: Mathematical Language**

1. I didn’t use any math language correctly.
2. I used a couple of basic math words or notations accurately.
3. I went beyond basic math language and used the language accurately.
4. I relied heavily on sophisticated math language to communicate my solution.

**C 2: Mathematical Representation**

1. I didn’t use any representation.
2. I attempted to use appropriate representation.
3. I used appropriate math representation accurately.
4. I relied heavily on sophisticated graphs, tables, charts and/or drawings to communicate my solution.

**C 3: Presentation**

1. My response is unclear.
2. My response contains some clear parts.
3. If someone else read my response they would have to fill in some details to understand my solution.

**Whence the Good Problems?**

The teachers bring a new set of possibilities from their folders—some garnered from textbooks, some from conversations with colleagues, some adapted from older problems. One has worked with her school’s home economics teachers on calculating the cost of baking a turkey; another asked students to calculate the area of a leaf; another challenged a class to distribute a sports advertising budget among different media. Jean Sequeira, of Lamoille Union High
Vermont’s Pioneering Portfolio Process

Why all the fuss about what goes into Vermont portfolios? The reason comes from a pioneering move this state’s education department took in 1988—not to rely on conventional standardized tests to assess its students’ progress, but to assess the writing and math skills of all the state’s fourth and eighth graders using portfolios as a key measure. A 90-minute uniform test (which includes both multiple-choice and “open-ended” questions) is augmented with student portfolios showing either their “best pieces” in a specific area or a combination of best pieces and other work completed throughout the year.

In this small and relatively unpopulated state, it seemed an innovative yet logical path to more individual and authentic assessment. And because Vermont was the first state to measure student progress statewide this way, educators throughout the country are watching closely.

The first task was formidable: If portfolios were to be reliable and valid measures of achievement, the state would have to develop in its teacher corps across-the-board professional agreement—on what to include as evidence, on how to score it, and on how to prepare students to produce it. First, a cadre of master teachers met to hash out the answers to those tough questions. They devised rubrics in math and in writing; and then their colleagues—like those in this Randolph meeting—were recruited to try them out. Over the course of the last few years, in 17 small regional networks that covered the state, every fourth- and eighth-grade Vermont teacher has received training on the standards for student achievement. The focus now is to help them improve their teaching to achieve higher levels of classroom performance; at the same time, they are reaching out to their fellows in other grades enlisting their help in coaching students steadily toward those demonstrable goals.

One key question in the minds of critics is what use will be made of the data that results from portfolio assessments. It is unfair, many educators believe, to publish scores as a means of comparing school to school and district to district, as has occasionally been the case in the last year. Instead, they suggest, the strengths and weaknesses that portfolios reveal should be shared among teachers and students, parents and community, and ultimately used to inform and improve the school’s curriculum and teaching practices.

School in Hyde Park, Vermont, passes around a problem about train shunting, which required that two railroad cars avoid collision by entering and leaving a siding with the fewest possible engine reversals.

“A good problem for one person may not be good for another,” comments Jill Rosenblum, the state’s math portfolio coordinator, who is facilitating this workshop. “Some people are interested in football; others aren’t. Some people are perplexed by things that others understand readily.” The best problems challenge students with differing levels of ability and achievement. Solving the Tower of Hanoi with 3 disks, for instance, involves the same quality of challenge and thoughtfulness that solving it with 20 does for a more experienced student.

One teacher praises the University of Chicago math series developed for K-3 and 7-8. “It’s a tremendously ambitious approach,” she says, “and I had to prepare the parents and the school board for it. But my kids have met the challenge. And the problems fit right in as portfolio pieces; I never have to worry about coming up with good ones.” But several other teachers at the table observe that they have given up textbooks altogether, gathering their own resources from several sources instead. “It took the parents a while to get used to the idea,” Mark Johnson says. “I told them I had a whole stack of textbooks and anyone was welcome to take one home.”

“We’re reluctant to get into any one set of standardized tasks that we give out for everyone to use,” observes Rosenblum. “That’s partly so teachers and kids can generate mathematical problems out of the kids’ own interests. But more important, when teachers have to evaluate problems themselves, they learn much more about what makes a good problem. They’re actively learning something from the inside out, instead of just passively following another textbook method they may not understand. It’s harder at first, but the long-term results are better.”

“These problems are really driving instruction in my school—not just in the fourth and eighth grade, but in the other grades, too.”

Teachers Revising the Rules

As they disband, I ask these teachers whether their new assessment method has changed the way their students perceive math. Yes, yes, yes, they reply as a chorus. “All this is hard to learn and hard to do,” one
Panel Assessments: Unlocking Math Exams*

by Nancy M. Haas
and Sam LoPresto

Mathematics is a subject noted for “drill and practice,” where learning is often equated with the quantity of problem solving. Over the past three years, teachers at Holt High School in Holt, Michigan, have tried to shift our emphasis toward conceptual understandings and the active demonstration of learning. In particular, we have had great success with a panel performance assessment in Algebra II and Precalculus.

We began changing our curriculum in response to a new conceptual change model, developed by the Holmes Group at Michigan State University. The model requires us to learn whether students understand the concepts behind the subjects being taught rather than simply being able to recount facts and knowledge. Central principles of the model include: teach for understanding; organize the school and its classrooms as a community of learning; hold ambitious learning goals; teach adults as well as children; make reflection and inquiry central features; and invent a new organizational structure for the school.

One of our teachers of Algebra II and Precalculus, Michael Lehman, began to alter his teaching and testing style to reflect these principles. Because he realized that multiple-choice or problem/answer exams would not effectively assess learning based on the conceptual change model, Lehman developed a panel performance assessment process.

How Panel Assessment Works

For their midterm and semester exams, students in Algebra II and Precalculus discuss before a panel of three judges the solutions to complex problems. Several days prior to the exam date, students receive six problems based on the main concepts discussed during the semester. The problems require some computation along with opportunities to make judgments based on the results of the computations.

Students work in cooperative learning groups of three or four to solve the problems and practice presenting their understanding of each one. They use their books and notes and ask questions of the teacher or other classmates as they feel the need. They often meet outside of class time as a group to continue their preparations. Parents assist with practice sessions before exam day arrives.

To assess the students, we recruit panels of judges from the business community, from Michigan State University (with which we have a strong partnership because of our role as a professional development school), from our school board, and from among parents and other district administrators. Educators from nearby school districts who are curious to know more about the panel assessment process also sometimes serve as judges.

Each panel consists of one “math expert” (usually a math teacher, a university math faculty member, or a math-related professional) and two “laypersons” (judges who may not necessarily be well versed in either algebra or precalc-

Having laypersons on the panel allows us to check for understanding in everyday language; students need to articulate and explain the math concepts to those who may not have the background for understanding them well. Prior to the exam, the panel judges receive the complete set of problems, the answers, and the scoring procedures.

On the day of the exams, student groups go before the panels of judges as a team. Currently, the assessments take place in the classroom and the school library, where tables are arranged so that no two groups of students are able to see each other. Each panel assesses only one team for Algebra II and one team for Precalculus.

The judges randomly select one of the problems and ask a team member to explain the problem. The students must be comfortable with all of the problems, as each will be asked to answer one of the six individually. After the judges hear a student discuss the problem, they may probe more deeply, or they may open the discussion to other students on the team.

After the panel assessment, each judge rates the individual students on both their understanding of the mathematics concepts described and on their ability to present what they know. Each student receives a letter grade based on a point system, as well as written descriptions of his or her abilities to communicate, to make sense of the problem, and to relate it to other “real-world” examples. This form of assessment gives students a much broader picture of their mathematical understanding, as well as of their teamwork and communication skills.

Students wait anxiously for the results of their panel assessment. A student who has never scored above a C on a traditional math test and who had failed more tests than he had passed told the math teacher he was grateful that the judges took the time to ask questions; he knew the information, but had trouble finding the right words. When this student found out that he had received a B on the exam, he jumped two feet off the ground and ran down the hall shouting the results!

The Payoffs

Organizing panel performance assessment does require a considerable amount of time. The time, however, is a trade-off against the time spent grading traditional exams, as the grading is completed on the spot. More important, this form of assessment is a rich supplement to the written methods used in the math classes.

The panel assessment process offers students an alternative to the traditional paper-and-pencil tests. Many students do not memorize information well, or they freeze when faced with a sheet full of complex computational exercises. During the performance assessment, these students have the opportunity to seek clarification from the judges and to express themselves orally as well as in writing or with other visual aids. Students gain a confidence that some of them have seldom experienced.

Performance assessment also provides students with a more meaningful way to demonstrate what they have learned. Students organize themselves and focus on what they need to know or do to successfully demonstrate their learning. This relevant, real-world assessment empowers students to construct knowledge and apply complex thinking skills.

Several teachers in other content areas such as global studies and American studies are now exploring using a similar technique in their classes. As we move forward to make performance assessment the norm rather than the exception at Holt High School, we anticipate that assessment, instruction (driven by performance assessment), and student achievement will improve, pointing students toward future success in their learning and work experiences.

1See the Holmes Group, (1990), Tomorrow's Schools: Principles for the Design of Professional Development Schools (East Lansing, Mich.: The Holmes Group).

2The math panel assessment and several other innovations are detailed in Alternative Assessment: Emerging Theory and Practice at Holt High School, B. Katney, ed., (Holt, Mich.: Holt Public Schools, 1993). Copies may be purchased for $12 each by writing to Nancy Haas at the address below.

For more information contact:

Nancy M. Haas is Assistant Principal, Holt High School, 1784 Aurelius Rd., Holt, MI 48842.
Sam LoPresto is Assistant Superintendent, Holt Public Schools, 4610 Spahr Ave., Holt, MI 48842.
Changing Minds is published by the Educational Extension Service, a component of the Michigan Partnership for New Education. The Partnership brings together schools, families and communities, business and industry, universities, government; and other concerned citizens to discover and demonstrate new educational approaches that enable both students and educators to thrive in a complex democracy and a global economy. Partnership member institutions work intensively with a growing number of Michigan schools and communities to improve learning, both in school and out, for young people and for the educators and other professionals who nurture their growth. These Professional Development Schools (see list) and the emerging Local Area Partnerships that will support them develop new insights, ideas, know-how, and tools for potential use by other Michigan schools and communities.

The function of the Educational Extension Service (EES) is to communicate this practical, research-based knowledge to schools and communities worldwide. To do so, the EES collaborates with intermediate school districts, professional associations, colleges and universities, consortia, and other organizations that provide information, professional development, and assistance to local educators, youth service professionals, family and community members, business people, and other citizens committed to improved learning and development for the state’s young people. In addition to these human and organizational networks, the EES will also use two-way television, computer networks, and other technological networks to support the communication of knowledge and know-how statewide. Thus, EES is not a new, free-standing organization but a network of networks. Through a variety of projects, the EES is now working with over sixty partner organizations. Because the list of our partners is large, growing, and changing, it is not possible to recognize each by name. You know who you are, and we are honored to have you as colleagues.

The EES also creates products for dissemination through the network of networks, including Changing Minds. Each issue of Changing Minds focuses in some depth on a difficult problem or topic in the renewal of teaching, learning, and the organizations that support these functions. Changing Minds includes both summaries or distillations of research and descriptions of ongoing innovation efforts, principally within Michigan but also in other parts of the country. Our goal is to blend research with thoughtful innovative practice, thus contributing to the continuous renewal of education in Michigan.

Professional Development Schools

Central Michigan University
Mt. Pleasant
Mary McGuire Elementary School

Eastern Michigan University/
Mary Grove College/Oakland
University/University of Detroit-
Mercy/University of Michigan-
Dearborn/Wayne State University

Detroit
Katherine B. White Elementary
School
Emma Stark Hampton Elementary
School

Grand Valley State University
Michigan State University
Muscogon Heights
Muscogon Heights High School
Michigan State University/
Grand Valley State University

Muscogon
Nelson Elementary School

Lake Superior State University/
Michigan State University
Brinley
Brinley K-12 School

Michigan State University
East Lansing
Spartan Village Elementary School
Flint
Brownell Elementary School
Gundy Elementary School
Holmes Middle School
Northwestern High School

Holt
Elliott Elementary School
Holt Jr. High School
Holt High School

Lansing
Averill Elementary School
Kendron Elementary School
Otto Middle School
Eastern High School

Northern Michigan University
Eben Junction
Superior Central Elementary

Oakland University
Pontiac
Longfellow Elementary School
Pontiac Central High School

University of Michigan
Ann Arbor
Carpenter Elementary School
Ypsilanti
Edmonson Middle School

Western Michigan University/
Michigan State University
Battle Creek LaGrange
Prairieview Elementary School
Michigan State University/
Western Michigan University
Battle Creek
Central High School

William Diehl
203 Erickson

252 Erickson Hall
Michigan State University
East Lansing, Michigan 48824-1034

Non-Profit Org.
U.S. Postage
PAID
E. Lansing, MI
Permit No. 21