What's not on the test? Teachers, students, and parents are drawing new attention to the vital skills and habits that most state tests ignore—and asking for more and richer ways to show what they have learned.

BY KATHLEEN CUSHMAN

A GROUP OF NEW JERSEY fourth-graders spreads a map on the floor and calculates with a bar scale how far they must travel to reach Death Valley. They will use the information in the making of an animated film on national parks, including detailed material on their ecology, geography, and cultures.

At a large suburban high school near San Francisco, students spend weeks researching the school facilities crisis that faces not only their own area but the whole state of California. The project concludes with a trip to Sacramento, where students lobby legislators to enact solutions they propose.

In a year-long project required before graduation, seniors outside Seattle each choose a subject of intense personal interest, from observing gorilla communities to sound engineering technology. Guided by an outside mentor, they research their topics; apply what they learn by creating a product, service, system, or event; then defend and reflect on their learning in a public presentation.

What do we hope students learn from projects like these? How good is the evidence that they are learning it? The answers are vital—because in the energy, engagement, and hard work of these students and many more like them, we can find clues to knowledge that remains unrecognized by the assessments most of the nation's districts and states currently use.

Teachers in Essential schools around the country are quick to reply to the first question. Using words like critical thinking, problem-solving, communication, teamwork, and persistence, they describe “habits of mind” that cut across narrowly academic disciplines. Though students must know and use subject-area content to carry out these tasks, their success depends even more on competency in these broader, cross-cutting skills.

But to the second question—how good is the evidence of student learning?—the answer is far less clear. What students learn through such thoughtful work does not usually show up on the kind of large-scale standardized test now common in virtually every district and state.

It's Not on the Test

In the early days of the standards movement, states typically promised to hold schools accountable for student achievement in a number of ways. But because it can seem simpler and cheaper to use test scores alone, most states are now phasing in high-stakes standardized tests as the sole indicator of whether students and schools are up to snuff.

For students denied promotion or graduation on the basis of those tests—as well as for teachers who
## Which Assessment Works Best? A Matching Test for Educators

Match the statement on the left with the appropriate assessment method from the list at right. (Some methods may apply to more than one item.)

<table>
<thead>
<tr>
<th>Statement</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I want to figure out how to improve my teaching in my classroom.</td>
<td>A. I analyze scores of my students on a standardized test that ranks students on a curve in comparison to all other students in the state or nation who took the same test (a “norm-referenced” test).</td>
</tr>
<tr>
<td>2. I want to figure out how to revise my classroom curriculum.</td>
<td>B. I analyze scores of my students on a standardized test that measures whether a student has reached a particular level (such as “basic” or “proficient”) in relation to a particular set of state or national standards (a “standards-referenced” test).</td>
</tr>
<tr>
<td>3. I want to figure out how to place students in different levels of instruction.</td>
<td>C. I ask my students to show me that they can “do” the subject I have taught, by using the knowledge and skill in carrying out a complex task to the standards set in my classroom or school.</td>
</tr>
<tr>
<td>4. I want to assess students’ “habits of mind” (such as evaluating evidence or making connections).</td>
<td>D. I develop and give my students a written test or quiz (multiple-choice, short-answer, or essay) that draws on the specific knowledge and skills I have been teaching.</td>
</tr>
<tr>
<td>5. I want to assess the basic skills of literacy or numeracy.</td>
<td>E. I solicit my students’ thoughts through an interview, survey, or written reflection.</td>
</tr>
<tr>
<td>6. I want to assess what content knowledge students have in a specific area.</td>
<td>F. I analyze student work using rubrics that describe performance at various levels.</td>
</tr>
<tr>
<td>7. I want to figure out how to improve our school’s program or set new school-wide goals.</td>
<td>G. I observe students at work on a project.</td>
</tr>
<tr>
<td>8. I want to check whether all students have the opportunity to learn at challenging levels.</td>
<td></td>
</tr>
</tbody>
</table>

---

可以丢掉他们的工作，而学校可能面临关闭和“重组”的风险—这一决定有严重的后果。随着它变得越来越难证明任何一项工作，使时间远离测试准备，许多教师已经放弃了项目工作。相反，他们将课程严格限制在测试—往往使用练习题和练习册出售给测试公司的自身。

惊警的是，那些将丰富而个性化任务和学术领域任务日益意识到需要评估、记录和报告给公众学生所学知识。他们也更加精明在为一个混合有意义的措施来展示显著学生学习。

### 什么知识持久？

在学生提出一个关于潮池的令人印象深刻的生物学项目给一个沿海加利福尼亚州的听众后，对卡特・詹梅茨，谁是西部评估合作的西方评估实验室在加利福尼亚州旧金山分部的主管，而言的时刻到来。“所以这些孩子知道他们去蒙大拿州时会做什么？”一位母亲在讨论后问。

詹梅茨说，这个问题去到一个在信息过载和选择性标准化测试中面对教师的困境。她去到”能力计数“，一个由coalition of essential schools和Jobs for the Future，一个专注于社区连接学习的国家组织，主办的三月2000年会，在旧金山召开。会议呼吁全国范围的中学教师。

詹梅茨说，我们应识别这些项目中的持久知识，”她声明去到"能力计数"，一个由coalition of essential schools和Jobs for the Future，一个专注于社区连接学习的国家组织，主办的三月2000年会，在旧金山召开。会议呼吁全国范围的中学教师。

詹梅茨说，我们应识别这些项目中的持久知识，”她声明去到"能力计数"，一个由coalition of essential schools和Jobs for the Future，一个专注于社区连接学习的国家组织，主办的三月2000年会，在旧金山召开。会议呼吁全国范围的中学教师。

詹梅茨说，我们应识别这些项目中的持久知识，”她声明去到"能力计数"，一个由coalition of essential schools和Jobs for the Future，一个专注于社区连接学习的国家组织，主办的三月2000年会，在旧金山召开。会议呼吁全国范围的中学教师。
scend particular subject areas.

The tide pool project, they agreed, provides students specific knowledge about marine biology. But it also gives them vital training in scientific investigation, data analysis, presentation methods, and even teamwork—all of which could apply just as easily to the somewhat different environmental questions that might come up in a classroom in Montana.

Building this kind of enduring knowledge requires skills and habits of mind difficult to measure in standardized assessments, these teachers noted. But to capture evidence of such learning, they needed new ways to observe, assess, and document it.

At Henry M. Jackson High School north of Seattle, for example, Judith Gray wanted her ninth-grade science students to understand not only the theories that explain the physical world, but also how scientific knowledge builds as people in each era seek out and correct fallacies in prevailing theories. Drawing on a suggestion in Neil Postman’s 1996 book *The End of Education*, Gray designed a performance task for her final take-home exam.

“Describe five significant scientific errors scientists have made, as well as why these were errors, who made them, and who was mainly responsible for correcting them,” she told her students. For extra credit, they could point out an error made in correcting the error, suggest a possible error in current thinking about science, or describe a possible error in one of their own firm beliefs.

As students explained errors and corrections from Galileo to Hubble, Darwin to Einstein, continental drift to the ionosphere of Mars, Gray got a good sense of the science each understood, and of what it would take to move their understanding along. And students learned not only the scientific facts, but the enormous impact that individuals can wield when they question—in any field of knowledge—the assumptions that “everybody knows.”

**Accountability Is Local**

Once teachers like Gray have clearly defined the mix of content knowledge and habits of mind they most want students to acquire, they can arrange both instruction and assessment around those priorities. But most states also make students take standardized tests in an array of subjects. When their high stakes narrow curriculum to what will be tested, they can throw a wrench into teachers’ efforts to nurture Essential school habits of mind.

With this in mind, CES’s regional Center in Maine is working with four Essential high schools to dovetail state tests with a ground-breaking system of “learner-centered accountability.”

Maine has its own standardized assessments to test its Learning Results, which consist of broad outcomes in the form of “guiding principles” (valuing habits of mind) and content and performance standards in eight areas. But local assessments count for far more than state tests in decisions about promotion and graduation. In Maine’s tiny districts, the school itself typically remains the most important arbiter of student progress.

“Our Learner-Centered Accountability project asks each school to identify which of the Learning Results they consider most vital,” says David Ruff, who directs the CES Center at the Southern Maine Partnership. “Then we help them create assessments of students’ progress toward meeting those.”

Teachers of the same subject at a particular grade level (like high school U.S. history) agree on a set of vital outcomes for all students. Then they devise common assessments to ensure that all students master certain basic material. Individual classroom teachers may tailor additional assessments to suit their own course emphases.

At yet another level, schools use portfolios or exhibitions to assess broader habits and skills, such as communication. These may be linked to a course or not, as with a senior project. (See sidebar, page 4.)

“The state tests serve as a validity check,” Ruff says. “If we believe our students can read well, for instance, and state tests tell us they can’t, we check out where the prob-

---

**Valid and Reliable? Test Your Own Task**

Teachers who like to use activities or projects to bring instruction to life may also assume that such activities make valid and reliable assessments of what a student understands. Not necessarily, warns Grant Wiggins in his 1998 book, *Educative Assessment*—but it’s simple to check, using these two questions:

1. Could the student do well at the task for reasons that have little to do with the desired understanding or skill being tested?
2. Could the student do poorly at the task for reasons that have little to do with the desired understanding or skill?

If either answer is yes, Wiggins says, either the task is an invalid measure or the evidence it produces will be insufficient or misleading. Instead, before designing an assessment the teacher should state precisely:

- What new skills and knowledge the student should be able to demonstrate.
- What constitutes the necessary evidence of that learning.
- Where one must look for that evidence.
bad day. This puts the focus of accountability where it belongs: in the school and community."

Standards of the Real World
As Maine's example shows, part of any accountability effort involves teachers and communities deciding which "competencies"—among the huge number outlined in different sets of standards—are most important to a student's future success. Many teachers at the CES-JFF conference brought examples of holding students not just to strictly academic standards, but to professional criteria used in the real world.

"The more meaningful the task in the community," noted Anne Purdy, who coordinates senior internships at Central Park East Secondary School in New York City, "the more likely it is that legitimate external partners can validate the learning involved."

At her school, older students complete 100-hour internships that culminate in a "service learning portfolio" required for graduation. If they choose to count this as one of five "major" portfolios they defend in juried presentations, students also develop a substantial project deriving from their workplace experience.

Working in the community relations office of Mount Sinai Hospital, for instance, senior Lynette Gonzalez created a bilingual newsletter aimed at reducing prenatal and early childhood exposure to toxic substances among community residents. Lynette researched and wrote informative articles on pest management, lead, head lice treatments, and PCB contaminants in fish. Then she laid out the newsletter, using desktop publishing techniques and meeting deadlines. "It felt like an endless process," she wrote in one journal entry. But by her last day on the job she and her supervisor both assessed the newsletter as a success.

Another student, Naimah Martin, worked with the Educational Video Center to produce a documentary on teenagers coping with death. Her portfolio included detailed assessments of not only the sound production, research, interviewing, and directing skills she mastered along the way, but also Naimah's skills in collaboration, listening, and concentration. Finally, because Central Park East assesses all graduation portfolio presentations according to schoolwide cross-cutting standards, Naimah's project was scrutinized for its viewpoint, connections, evidence, voice, and conventions.

The presence of adults who can mentor students in real-world standards for high-quality work makes an enormous difference in projects like these. Coached by graduate students from the University of California, for example, students at Oakland Technical High School carried out a six-week urban design project in a neighborhood near their school. (See sidebar, page 5.)

They began by mapping their community, getting familiar with...
planning and design terminology and exploring local history by interviewing residents. Then they created a detailed physical survey of the area surrounding a local subway stop, and teams of students assessed the neighborhood’s politics and economy. After presenting their site analyses to each other, they held a design “charrette,” pooling their observations and ideas to create a viable plan complete with a financial feasibility analysis. Ultimately, the students presented their plan to a jury of city officials and planning professionals.

In Fort Lauderdale, Florida, Stranahan High School runs a six-week summer “science research institute” in which students conduct field studies on actual environmental projects from fire management to wetlands regeneration. Mentored by environmental scientists, they obtain samples, use professional methods to analyze field and lab data, summarize current research in writing, and present their work before local county audiences. The program has won such praise that Stranahan is seeking ways to conduct others like it during the year.

Assessing Rigor and Habits

After spending two days in critique of projects like these, the educators gathered at the CES-JFF conference agreed on several elements crucial to ensuring that student tasks and assessments expect both rigor and the habits of lifelong learners:

1. Tasks should involve research that expands a student’s ability to ask questions and seek answers.
2. Throughout the task, students should have adult coaches who can help them understand what high quality looks like in the areas the task involves.
3. Tasks should build in student revision, following critical feedback from key audiences, and continuing until the work meets the appropriate expectations.
4. Some means of follow-up should seek to determine what learning endures after the task is finished.

In this 90-minute class from a six-week urban design project focusing on their own Oakland, California neighborhood, for example, high school students are learning economics principles through action. The lesson aims:

- To provide an overview of economic disparities and their impacts on neighborhoods and communities.
- To explain the concept of the income multiplier, and how it affects a local economy.
- To consider different interests involved in the development process.
- To discuss how a bank determines risk.
- To investigate causes, effects, and potential solutions.

By using field studies on actual environments, students can begin to understand how goods and services are distributed in distressed communities, including how retail and commercial clients—such as the Oakland Raiders or the downtown ice rink—make decisions on where they will locate. To link the multiplier effect of circulating capital locally, they use diagrams of money flow (to employees, to sales tax, to national headquarters, etc.), and explore the trade-offs between national chain stores and locally owned businesses. To prompt student inquiry and assess their understanding of these economic principles, teachers note their answers to questions like these:

- Where do most people in your neighborhood buy groceries?
- What are the benefits of having a supermarket in your neighborhood?
- What does the money you spend at a supermarket?
- How could you get those benefits, but keep more money in the local economy?

For more information about the Urban Plan project curriculum, contact Patricia Clark by e-mail at pcclark@acoe.k12.ca.us.

In Search of the Right Mix

Qualities like self-reflectiveness, persistence, creativity, collaboration, and learning by discovery are difficult to teach and assess by conventional methods, these educators worried. Equally tricky are the knowledge and ways of thinking that do not emerge when a teacher has one eye on a standardized test. “You can test a student’s recall of the names of marine organisms, for example,” says Kathy Simon, CES’s director of research and professional development. “But it’s rare for a standardized test to get at what happens over a whole life cycle and in an entire ecosystem. What are the implications if the starfish die off?”
One Community Action Research Project and the Standards It Met

Eleventh and twelfth grade students in “Academy X,” a leadership and humanities academy at Sir Francis Drake High School in suburban Marin County, California spent nine weeks researching the school facilities crisis that faces not only their own area but the whole state.

Working in groups, the students researched the facilities problem by meeting with school officials and state policy-makers and visiting schools. Next, groups prepared a variety of tools (web sites, videos, written proposals) to publicize the problem and their proposed solutions. The project concluded with the students arranging a trip to Sacramento (complete with press coverage), where they lobbied legislators to enact their suggestions.

“The essential question for Academy X’s year of learning, was: What do I need to know in order to effect positive change in my community, my school, and myself?” says Bob Lenz, one of a team of social studies, English, and Workplace Learning teachers who led the project, known as Community Action Research Go! (or CARGO). “We wanted students to have a real world context for their study of government, economics, and oral and written communication.”

Some of the project’s other goals for student learning:

- An understanding of the legislative process by trying to access the California State Legislature.
- An application of the economic principles of scarcity, trade-offs, opportunity costs, investment, economic growth and long run vs. short run decisions through the defense of their proposals to the legislature.
- A context for reflection on the power and pitfalls of collaborative problem solving and community action.
- Persuasive writing and speaking to a real audience about an issue the student finds relevant or meaningful.
- An introduction to the field-study and problem-based learning models students would soon apply during individual internship experiences.
- A meaningful or relevant field trip to the state capitol.

Students were expected to finish their research and “deliverables” at a professional standard, and their products were assessed using rubrics the students devised. The following individual and group tasks were also assessed:

Individual:

1. A speech to inform on a community issue with governmental or economic implications. (Rubric for oral presentation.)
2. A brief research analysis of individual research on group topic. (Rubric for an investigative report.)
3. A speech to persuade. (Rubric for oral presentation.)
4. A journal of activities and reflections as work on the project progressed. (Feedback using rubrics for group skills and for time and task management.)

Group:

1. A summary of their research. (Exhibition rubric assessing action research, content, defense of argument, and oral presentation skills.)
2. The product or arrangements they agreed to make. (Rubrics designed by the students.)

Students spent two 90-minute periods weekly for nine weeks (a total of 14 hours) on the project, as well as out-of-class time as necessary. In the process, Lenz believes, they met both broad and specific standards set by their state and district, having to do with:

- Understanding and application of economic principles, concepts, terms, and reasoning.
- The roles and responsibilities of the three branches of government under the U.S. and state constitutions.
- Effective written and spoken communication to inform and persuade.
- Reading, viewing, and analysis of material from a variety of disciplines.
- The use of technology to access information.
- Analysis and problem solving of current issues from historical, political, and economic perspectives.
- Post-secondary and workplace transition skills and knowledge.
- Community, social, civic or cultural participation.

In the student’s own words, the criteria for success were considerably more vivid and concrete:

- Did we make it to Sacramento?
- When we were there, did we meet with key decision-makers?
- Did those decision-makers take our research, conclusions, and recommendations seriously?
- Did we know what we were talking about?
- Did the press cover our visit and our recommendations?
- Did we effect positive change?

“This project could be scaled down, using only the Internet and the phone for research, desktop publishing for the report and a trip to the local government,” Lenz points out. “Or it could be scaled up—perhaps a trip to Washington D.C.” He has compiled the details in a CD-ROM-based portfolio complete with sample student work and rubrics, which he presented to teachers from around the country at “Competencies That Count,” a conference sponsored in March 2000 by the Coalition of Essential Schools and Jobs for the Future, a Boston-based organization focused on community-connected learning. For more information about the project, contact Bob Lenz at blenz@marin.k12.ca.us.
Schools cannot rely on state tests to assess such crucial elements of student learning, the teachers gathered in Oakland agreed. Instead, “We need a mix of pedagogical styles for kids to learn the range of knowledge, skills, and ways of thinking we’re looking for,” Simon asserted, “and a mix of measures to capture them.”

Deborah Meier, Vice-Chair of the Coalition of Essential Schools and director of the Mission Hill School in Boston is even more emphatic. Assessing the learning of children, she says, requires “a wide range of tools, always administered individually, alongside samples of real work, plus opportunities to check them out with the child and those who know him or her best.” “If we’re talking about assessment for the purpose of ‘accountability’ or certifying students as properly ‘ready,’” she adds, “that’s another story.” The driver’s test provides an admirable standardized performance assessment in this last context, she suggests, conceding that “few intellectual skills lend

---

The Senior Project: Demonstrating Academics Alongside Life Skills

Schools around the country have turned to Senior Projects as a way to synthesize and demonstrate a student’s intellectual as well as life skills. Typically, such projects arise out of students’ individual passions or interests and are mentored by an outside expert in the field. At Henry M. Jackson High School outside Seattle, students spend an entire year pulling together their independent projects. A few of some 350 projects presented at Jackson High School in 2000:

- A senior girl explored the demands of marathon running, and trained to run a marathon herself.
- A boy with a passion for electronics built his own sound board, then served as recording engineer for a local band’s CD.
- Fascinated by the interactions of gorilla communities, one girl observed a gorilla colony weekly at the local zoo and conducted e-mail interviews with Jane Goodall.

Senior Project is a year-long class at Jackson, and the work is assessed throughout, with rubrics for each stage. The May presentation culminates the process, with students describing to a jury of Senior Project teachers-assessors their guiding questions, research, project activities, and resulting learning. Guidelines for each stage are explicit, as follows:

**Preliminary Steps (September through mid-November)**

1. Review your life goals and aspirations. Identify areas of interest and subjects that interest you.
2. Make initial primary source contacts, and investigate other likely sources.
3. Begin a journal that records activities, observations, and responses related to your learning.
4. Locate and confirm a mentor or instructor and submit the Mentor/Instructor Correspondence Form and the Parent Approval Form.
5. Meet with your mentor to complete the Senior Project Proposal Form and Action Plan, then submit it to your Senior Project teacher for approval.
6. Revise the proposal if necessary.

---

Project Phase I (Annotated Bibliography and Project Activities) (mid-November through early April)

**Step I. Preliminary Investigation**

1. Investigate what you need to know in order to accomplish your learning goals.
2. Investigate where you can find that information.

**Step II. Guiding Questions**

1. Develop questions to guide your information-gathering process.
2. Use your guiding questions to gather information from interviews and other sources.

**Step III. Annotations and Citations**

1. Write an annotation for each source that summarizes the answers to your guiding questions.
2. Evaluate the credibility and usefulness of each source.
3. Cite your sources correctly in MLA style.

**Step IV. Project Activities and Documentation**

1. Maintain a regular, updated journal recording your progress.
3. Meet regularly with your mentor for feedback and advice.
4. Submit your documentation of progress.
5. Complete and submit your project for evaluation.

Presentation Phase (mid-April to mid-May)

1. Prepare an abstract that details the main points of your project phase.
2. Prepare a visual or technological component for your presentation, if appropriate.
3. Prepare the entire oral presentation of your project.
4. Practice your presentation.
5. Make your presentation to the panel.

Reflective Phase (late May and early June)

1. Write a rough draft of your reflective paper.
2. Revise your paper after feedback and conferencing.
3. Proofread, edit, and polish your paper.
4. Submit final draft for evaluation.

---

Complete descriptions and rubrics for Jackson’s Senior Project process are available at: [www.everett.wednet.edu/schools/high/hfm/jackson/](http://www.everett.wednet.edu/schools/high/hfm/jackson/)
themselves to this format. Schools should leave the certifying to colleges or workplaces, Meier argues, and concentrate instead on meeting the learning needs of individual students.

Inventing New Methods

In fact, many interesting tries at assessing competencies that transcend academic areas have emerged for use in high schools, community colleges and workplaces, notes Lili Allen, a senior project manager at Jobs for the Future. She is assembling a "road map" describing these methodologies, including several in use at Essential Schools.

For example, transcripts developed by California's Transitions Project and by the New Hampshire Department of Education focus on documenting such competencies for post-secondary institutions and employers. And the Massachusetts Work-Based Learning Plan has tools to give feedback to students in programs that link academic and career-related classes with work-based learning.

Schools in Fort Worth, Texas have tried out the "applied learning standards" designed by the New Standards Project, a partnership between the National Center on Education and the Economy and the University of Pittsburgh. Part of a package of performance standards that includes English Language Arts, Mathematics, and Science, these define skills like problem-solving and communication, calling for their integration and assessment as part of products or performances associated with academic assignments.

Finally, New York City's small Essential high schools are among many who track data on how their graduates fare after high school, in what Theodore Sizer calls "the ultimate assessment." One can learn much, he points out, from students whose strengths do not show up on test scores, but who are doing well in college and life because of habits they learned in school.

What Role for State Tests?

More conventional standardized tests can provide interesting large-scale data about issues of equity and access, these teachers agreed. Some states, for example, test only literacy and math skills, to identify programs that fail to give students equal opportunities to learn. And teachers at the conference praised certain state tests for prompting higher-order thinking.

But all emphatically objected to systems that reward or punish students and schools on the basis of such test scores. Instead, they called for a balanced assessment system that could guide their teaching practice day to day, honoring the real learning needs of students:

"A baseball coach doesn't look only at win-loss records when he works with a team," said Kate Jamentz. "What would that tell him? To build his players' skills, he also needs to observe a lot of other things about what his players can and can't do. He needs to see what's getting in the way of their scoring, and work on those things until they're doing better."

It's the relationship between the "win-loss record" and the diagnostic information that schools most need today, she asserted. "Otherwise, we're like a hospital that has mortality statistics but no X-ray machines," she said. "We know who's not achieving, but we don't know what to do about it."