

HORACE

Performances and Exhibitions: The Demonstration of Mastery

A true test asks students to show what they know and can do, not to spout unrelated facts they have memorized the night before. Once we start measuring performance this way, change is swift to follow—in what we teach, in how we teach, and in our assumptions about why kids are in school at all.

WHAT DO WE WANT HIGH school students to learn? The most revealing answer can be had by looking at what we expect from them when their time is up. What students know and what they can do, after a course is completed or a high school career ended, is in many ways a reflection of what their schools have expected them to master.

For many, this means proving themselves through a series of mechanized hoops—machine-scanned textbook tests, achievement and competency tests. But for students in an increasing number of schools, the question of mastery is becoming at once more messy and more authentic. What can you *really* do? teachers uncomfortable with what conventional tests show are beginning to ask. What do you understand about how to get answers to hard questions? And in realistic contexts before mixed audiences of peers, teachers, and the community, students in many schools are showing us the answers, in the exhibitions that expose the very heart of what the Coalition of Essential Schools is trying to do.

"In its original form, the exhibition is the public expression by a student of real command over what she's learned," says CES Chairman TheodoreSizer. "It began in the eighteenth century, as the exit demonstration in New England academies and in colleges like

Harvard. The student was expected to perform, recite, dispute, and answer challenges in public session." If such a performance is well designed, Sizer points out, it elicits proof both of the student's understanding and of some imaginative capability—it serves at once as evaluative agent and expressive tool. "We expect people to show us and explain to us how they *use* content—it's more than mere memory," Sizer says. "It's the first real step towards coming up with some ideas of their own."

The concept of performance-based evaluation is nothing new, notes Grant Wiggins, who has been a consultant to CES on assessment issues; we see it every time someone presents a business proposal, performs in a recital, plays a ball game. But the exhibition is at least as much a teaching tool as an assessment method, Sizer points out, as much inspiration as measurement. "Giving kids a really good target is the best way to teach them," he says. "And if the goal is cast in an interesting way, you greatly increase the chances of their achieving it. When you can see the obvious exhilaration of the final act—as, for example, in *really using* a foreign language well—it's perceived quite differently from the usual test, which is secret and comes at you in a way you never see in other areas, with time constraints and machine grading."

Figure 6. Walden III's Rite of Passage Experience (ROPE)

All seniors must demonstrate mastery in fifteen areas of knowledge and competence by completing a **portfolio**, a **project**, and six other **presentations** before a ROPE committee consisting of staff members (including the student's home room teacher), a student from the grade below, and an adult from the community. *Nine of the presentations are based on the materials in the portfolio and the project; the remaining six presentations are developed especially for the presentation process.*

The Portfolio. The portfolio, developed during the first semester of the senior year, is intended to be "a reflection and analysis of the graduating senior's own life and times." Its requirements are:

1. A written autobiography, descriptive, introspective, and analytical. School records and other indicators of participation may be included.
2. A reflection on work, including an analysis of the significance of the work experiences for the graduating senior's life. A resume can be included.
3. Two letters of recommendation (at minimum) from any sources chosen by the student.
4. A reading record including a bibliography, annotated if desired, and two mini-book reports. Reading test scores may be included.
5. An essay on ethics exhibiting contemplation of the subject and describing the student's own ethical code.
6. An artistic product or written report on art and an essay on artistic standards for judging quality in a chosen area of art.
7. A written report analyzing mass media: who or what controls mass media, toward what ends, and with what effects. Evidence of experience with mass media may be included.
8. A written summary and evaluation of the student's course work in science/technology; a written description of a scientific experiment illustrating the application of the scientific method; an analytical essay (with examples) on social consequences of science and technology; and an essay on the nature and use of computers in modern society.

The Project. Every graduating senior must write a library research-based paper that analyzes an event, set of events, or theme in American history. A national comparative approach can be used in the analysis. The student must be prepared to field questions about both the paper and an overview of American history during the presentations, which are given in the second semester of the senior year.

The Presentations. Each of the above eight components of the portfolio, plus the project, must be presented orally and in writing to the ROPE committee.

Six additional oral presentations are also required. For these, however, no written reports or new products are required by the committee. Supporting documents or other forms of evidence may be used. Assessment of proficiency is based on the demonstration of knowledge and skills during the presentations in each of the following areas:

To graduate from Walden III you must complete a portfolio, a written project, and fifteen oral presentations before two teachers, a peer, and an outside adult.

1. *Mathematics knowledge and skills is demonstrated by a combination of course evaluations, test results, and work sheets presented before the committee, and by the ability competently to field mathematics questions asked during the demonstration.*

2. *Knowledge of American government should be demonstrated by discussion of the purpose of government; the individual's relation to the state; the ideals, functions, and problems of American political institutions; and selected*

contemporary issues and political events. Supporting materials can be used.

3. *The personal proficiency demonstration requires the student to think about an organize a presentation about the requirements of adult living in our society in terms of personal fulfillment, social skills, and practical competencies; and to discuss his or her own strengths and weaknesses in everyday living skills (health, home economics, mechanics, etc.) and interpersonal relations.*

4. *Knowledge of geography should be demonstrated in a presentation that covers the basic principles and questions of the discipline; identification of basic landforms, places, and names; and the scientific and social significance of geographical information.*

5. *Evidence of the graduating senior's successful completion of a physical challenge must be presented to the ROPE committee.*

6. *A demonstration of competency in English (written and spoken) is provided in virtually all the portfolio and project requirements. These, and any additional evidence the graduating senior may wish to present to the committee, fulfill the requirements of the presentation in the English competency area.*

The above is drawn from the 1984 student handbook, "Walden III's Rite of Passage Experience," by Thomas Feeney, a teacher at Walden III, an alternative public school in Racine, Wisconsin. Preliminary annotations are by Grant Wiggins.

knowledge, put things together, or go looking for facts when an answer is unknown.

In the end, students at schools like these are asked to show they know much more than those in conventional schools. Not only must they meet competency requirements in traditional areas and produce work that integrates material across the disciplines, but they must demonstrate their mastery of skills that will carry them into adult life. Central Park East, for one, expects proficiency in everything from personal economics to voter registration procedures and operating a computer.

The "Rite of Passage Experience," or ROPE, program at Walden III, which has been in place for over a decade, is a fully developed model of how such a requirement can function. Born from the Australian "walkabout" tradition in which a youngster must meet certain challenges to attain adulthood, ROPE is expressly designed "to evaluate students' readiness for life beyond high school." In order to graduate, each senior must satisfactorily present a written portfolio, a written project, and an oral demonstration before a committee made up of staff, another student, and an adult from the community. The fifteen required areas of mastery cover a wide range; students not only demonstrate competency in academic disciplines but submit other materials like a written autobiography, a reflection on work, an essay on ethics, and a report analyzing mass media. (See Figure 6.) The ROPE requirements

Central Park East expects proficiency in everything from personal economics to voter registration procedures and using a computer.

are very close to the graduation exhibitions in use at Rochester's School Without Walls, and to those being developed at Central Park East Secondary School.

Students follow a strict schedule in preparing their projects and portfolios, which must be ready for presentation by the end of first semester. At Walden III, ROPE committees schedule as many as five meetings with each senior during the second semester—enough so that if weaknesses are evident, the student has the chance to remedy them. The whole process is expensive in terms of teacher time, notes Thomas Feeney, who teaches a first-semester class preparing seniors to present themselves to their ROPE committees. A minimum of ten hours of teacher time is spent examining each student—time which must come either directly from courses or from conference periods. "To carry it off you need either a very committed school board or a very flexible curriculum, or both," Feeney says.

At Central Park East, the first eleventh-grade class is launched this year on a two-year "Senior Institute" designed to culminate in an exit exhibition similar to Walden III's. The program includes not only traditional course work but internships and lab work, seminars and independent study in collaboration with local universities, and summer projects. Because it is oriented toward mastery rather than accumulation of credits, the Institute is time-flexible; it may take more or less than two years depending on the student's needs and desires. And it is personal; each student is coached and advised throughout by a teacher responsible for no more than fifteen students and two academic courses.

Evaluation of exit exhibitions follows the same kind of procedures used in classroom performances—except that a student rarely gets to the exhibition stage unless he or she is actually ready to demonstrate mastery. If there are problems, the

"Whereas students can sometimes squeak through on regular competency tests, with ROPE we get smacked right in the face with what they don't know."

committee usually catches them in an early review. Still, if a majority of the group eventually so votes, the student will go back for as much more preparation as he or she needs until true mastery is achieved.

"In a way, as a method of evaluating what students know at the end of high school perhaps it's *too good*," says Thomas Feeney. "We substitute our own competency demonstrations for the standardized tests in the areas required by the district. But whereas students can sometimes squeak through on regular competency tests, with ROPE we get smacked right in the face with what they don't know." Feeney directly attributes changes in Walden's curriculum requirements to this fact. "We now require two quarters of geography, for instance, to graduate," he says. And ROPE results spark intense discussion among faculty, he says, as to how and when certain subjects should be presented.

That is probably the chief reason that exit exhibitions are such a difficult and controversial way for schools to show their commitment to active learning. If they start by defining clear standards for student mastery, change in every area of the curriculum is unavoidable. After all, one cannot expect some students to be held to a new graduation standard and others to the old way, some teachers to demand only skills in rote and recall and others to ask students to think their way through harder challenges. If no consensus exists on what students

CES is exploring the possibility of starting several new schools in which the entire curriculum would flow from the exit exhibition.

should know and be able to do when school is over, a school may be split at its very core. Exit exhibitions, then, can be diagnostic tool or catalyst; they cannot be a neutral, "safe" assessment measure.

Does this mean that exit exhibitions are only possible in smaller, alternative high schools with structures that invite such individual assessments? "You can't really do them unless you know the kids," Ted Sizer responds. But big schools should address that, he says, by

reducing their crushing teaching loads, personalizing and restructuring the curriculum so there is time for this kind of attention. And once exhibitions become a way of teaching, he says, teachers and students both are energized by the results. "To say 'Our kind of kids or teachers wouldn't do it' is just a self-fulfilling prophecy," he asserts.

For students, making exit exhibitions the end of their school careers also creates deep and encompassing changes. CPE's Senior Institute, for example, starts with making a post-graduation plan, which gives direction to all the student's efforts toward the diploma. This becomes a key part of the portfolio reviewed by the graduation committee, but it serves a more important role as well: it answers the question, "Why am I in high school?" Throughout the final years of high school, students must come to terms with the ways everything they do reflects back on

that question and shapes its answer.

The "atmosphere of unanxiously high expectation and trust" which is a common principle of Essential schools is a key element in putting the exit exhibition into place. Students get a clear message that high school is over only when they can demonstrate competence in areas that mean something to them and to the community. Whether that revolutionizing principle can be as effective in schools that start their performances at the classroom level instead of by defining graduation requirements is an open question. Ted Sizer believes not; the Coalition is exploring the possibility of starting several new schools in which the entire curriculum would flow from the culminating exhibition. In the meantime, teachers and students in Essential schools continue to grapple at every level with how to assess true thoughtfulness, real mastery, and a broad range of student capacities. □



The Coalition of Essential Schools
Box 1938
Brown University
Providence, Rhode Island 02912

Non-Profit Org.
U.S. Postage
PAID
Providence RI
Permit No. 202

Figure 1. Qualities of "Authentic Performances"

Structure and Logistics

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

Intellectual Design Features

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

Grading and Scoring Standards

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

Fairness and Equity

- Ferret out and identify (perhaps hidden) strengths.
- Strike a constantly examined balance between honoring achievement and native skill or fortunate prior training.
- Minimize needless, unfair, and demoralizing comparisons.
- Allow appropriate room for student learning styles, aptitudes, and interests.
- Are attempted by all students, with the test "scaffolded up," not "dumbed down," as necessary.
- Reverse typical test-design procedures. A model task is first specified; then, a fair and reliable plan for scoring is devised.

These are provided by Grant Wiggins, former director of research at CES; he gives credit to Ted Sizer, Art Powell, Fred Newmann, and Doug Archbald and to the work of Peter Elbow and Robert Glaser for some of these criteria.

If it were not for a bureaucracy of schooling deeply invested in easily generated outcomes, common sense might dictate the steps in which teachers assess students. Of course we want students who are curious, who know how to approach new problems, who use reading and writing across the disciplines as a natural part of that process, who are thoughtful, able, and active citizens. And to get them we would merely make those goals known from the start, test for them regularly, and correct a student's course when necessary.

What complicates matters is an approach to testing that originated in an era when it still seemed possible, and necessary, to impart to young scholars a set body of information. In addition, educational theorists believed that the way to learn things was to break them down into their smaller components. Testing reflected those assumptions: it was discipline-specific, content-driven, easily shaped into multiple-choice instruments of assessment. Even reading and writing were taught and assessed this way, broken down into discrete components that could be tested and taught separately. In the effort to achieve consistency and a uniform standard, "subjectivity" became a bad word; and in the push for scientifically accurate assessment no one acknowledged that even the choice and wording of items on standardized tests reflected biases as real as those of any classroom teacher.

In our information-loaded age, that system has lost whatever intellectual credit it may once have had. We can't know all "the facts" anyway; and even if it were possible, theorists now think students learn best when facts are sought in their context, not in arbitrary sequences. Instead, Essential schools aim for teaching students how to find out and critically evaluate the facts they need in a particular situation—the thoughtful habits of mind that are sometimes called the "higher

Most standardized tests, critics say, lack the subtlety and sophistication needed to test critical thinking skills.

literacies." Most standardized tests, critics say, lack the subtlety and sophistication needed to test such critical thinking skills.

But the old system of testing still has bureaucratic usefulness, and so students still hear its message clearly: they are in school to be sorted, ranked, selected. For better students, this can rob their studies of excitement or intellectual purpose. For the less advantaged, though, it implies a guaranteed level of failure—because norm-referenced standardized tests distribute students along a bell-shaped curve which can be predicted in advance. Such methods have changed testing, many argue, from a teaching and learning tool to the point where it serves only a social and political purpose.



HORACE is published five times yearly at Brown University by the Coalition of Essential Schools, Box 1938, Brown University, Providence, RI 02912. Subscription is free. Publication of HORACE is supported by a grant from the Rockefeller Brothers Fund.

Editor: Kathleen Cushman
Managing Editor: Susan Fisher

To change that pattern, the Coalition of Essential Schools asserts, we must change the very *reason* students go to school. This must begin, Sizer says, with a new expectation: that all students can use their minds well. New incentives are next: real mastery of things they want to do well. Finally, schools must provide new proving grounds where they can show off that mastery in positive, public, and personal ways. This last is known, in Essential schools, as the *exhibition*; both in theory and in practice, it is the cornerstone of what an Essential school is all about.

What Deserves a Diploma?

In an ideal Essential school, Ted Sizer believes, all decisions about a school's curriculum should flow from the devising of a culminating exhibition at graduation. Do we want graduates to be able to synthesize information from a variety of disciplines in a well-reasoned argument? Then design courses that give them regular practice in cross-disciplinary inquiry, and require a final project that shows they can do it. Do we want them to answer and ask questions on their feet, to work productively in groups? Then course work must consciously train them in these skills. Do we want them to judge the reasonableness of an answer, whether in mathematics or ethics, and to evaluate the quality of evidence? Then in every class give priority to such habits of mind over traditional coverage of content. Do we want active citizens who know their rights and ways to affect their own government? Then require courses that directly engage them in such matters.

Clearly, the structural choices that follow such an evaluation of a school's ends can be uncomfortably radical. They will affect every teacher and every student, at every level from the daily lesson plan to the final graduation hurdles. It is easy to see why so few in the

Coalition have put the "exit exhibition" first in their efforts to revise and restructure their schools. If one starts by defining graduation requirements in terms of demonstrations of mastery, it's difficult to proceed in cautious little steps.

Instead, most Essential schools have held off on developing a culminating performance before graduation. They prefer to develop, within individual or team-taught courses, a new style of assessing student progress that relies more on demonstration of thoughtful habits of mind and less on memorization of facts. These course-level exhibitions are referred to within the Coalition as "performances," to distinguish them from the graduation exhibition.

Just what do these performances look like? How are they graded, and what allowance is made for different levels of student ability? Without standard measures, how can a teacher reliably tell if the basic competencies are being mastered? Doesn't it cost a lot to do things this way, in teacher time and training, in administering and scoring? Don't we need to teach students to take the standardized tests that the real world judges them by? I asked these questions of Essential school theorists and teachers who have been trying performance-based assessment in their classrooms.

What Performances Look Like

At the classroom level, a performance is often as simple as a final essay that requires skills in inquiry and synthesis to answer what the Coalition calls "essential questions." [See HORACE, Volume 5, No. 5.] Or it might display student mastery in the form of a project, perhaps undertaken by a group. In some classes students prepare portfolios of their best work to submit for evaluation; in others, they present their work orally and answer questions on it before the class. Whatever its form, the performance must engage the student in real

intellectual work, not just memorization or recall. The "content" students master in the process is the means to an end, not the end itself.

Because Essential school teachers use such skills as part of their everyday commitment to "active learning," it can be hard to tell where performances start and regular classwork leaves off. And indeed, everyone agrees that performances do serve as a teaching tool as much as an assessment tool. But if we are to consider performances as an alternative to conventional testing, it is most useful to look at their evaluative purpose.

For example, at Springdale High School in Arkansas, humanities teacher Melinda Nickle and the other members of her teaching team devised a final exam that could be used for classes in inquiry and expression, literature and fine arts, social studies, and science. (See Figure 2.) Students are asked to link research materials across the disciplines in a thematic essay; later they participate in group evaluations of each others' papers. The format is aimed at many essential skills at once: an interdisciplinary approach, "student-as-worker," the development of critical thinking, and so on. It took a lot of time, Nickle says: at least two teachers had to evaluate

At the classroom level, an exhibition can be a final essay, a project, a portfolio, or an oral presentation.

each paper and then compare scores before a grade was put on them. But "all our students remarked on how much they learned from the performance," she says. "They spent hours working on it at a time of year when most students had already shut down

Figure 2. A Final Performance Across the Disciplines

Discuss behavior patterns as reflected in the insect world, in animals, in human beings, and in literature. Be sure to include references to your course work over the term in Inquiry and Expression, Literature and the Arts, Social Studies, and Science. This may include *Macbeth*, the drug prevention and communication workshop, Stephen Crane's poetry, "A Modest Proposal" and other essays you have studied, Mark Twain's fiction, and behaviors you have observed in our School-within-a-School. You may also add references to what you have read about in the news recently.

Procedure. *Day one of the exam:* You will be given four periods in which to brainstorm, make an outline, write a rough draft, and write a final copy in standard composition form. You will be graded not only on how well you assimilate the material but also how well you reflect the "student as worker" metaphor and how responsibly you act during the testing period.

Day two of the exam: You will assemble in villages of three, evaluate anonymous papers according to a set of criteria, and come to a consensus about a grade. Each paper will be evaluated by at least two groups and two instructors. Again, a part of your overall semester grade will have to do with how responsibly you act and how well you demonstrate the "student as worker" metaphor.

Thanks to Melinda Nickle at Springdale High School, Springdale, Arkansas.

due to the approaching vacation."

At Adelphi Academy in Brooklyn, New York, science teacher Chet Pielock and humanities teacher Loretta Brady ask students to form teams to investigate Latin America's problems of poverty and illiteracy, overcrowding, earthquakes, and political instability. (See Figure 3.) To answer some questions in the performance students must exhibit detailed geographic knowledge; to answer others, they must relate them to society and history. Interesting issues can arise from such work: How are "natural resources" regarded by different cultures? What happens when different cultures conflict over the value of a natural resource? How do natural resources function in human struggles for power?

The best performances and exhibitions are not merely projects aimed at motivating students; they evoke fundamental questions within a discipline. For a final exam in both history and English, for example, one teaching team has

students support or refute the statement, "What matters in history is not societies or events, but individuals." (See Figure 4.) Because it asks the essential question "What causes history?", such an inquiry can reflect not only how we see the past but how we think about the present and future. Next students are asked to evaluate their own essay along specific criteria, and then to relate an "English" essay on subjectivity in research writing to it as well—revealing the interdisciplinary connections between literature and history.

How to Grade a Performance

One of the reasons conventional tests hold such sway in schools, of course, is that they are easy to grade. Are teachers' evaluations of exhibitions and performances as objective and reliable as the multiple-choice and fill-in-the-blanks tests they replace?

"We can't evade these very technical questions of reliability and validity," Grant Wiggins said

Figure 3. A Final Performance Across the Disciplines

So you want to understand Latin America's problems: poverty and illiteracy, overcrowding, earthquakes, and political instability (that's right, war). Can Latin America overcome these problems? Is the United States helping Latin America all it can? What is behind these problems? One key topic we need to understand is the land itself. We will become three teams of experts exploring three key fields:

1. What forces have shaped the land?
2. What are the resources of the land?
3. Why do people live where they do on the land?

Each group has special requirements they need to fulfill, but in general, these are the expectations we have for you while you work and for the day of the final exhibition of your work:

"A" grade:

Everyone on team understands all the material well.
The team teaches the rest of the class effectively.
All diagrams and maps are effective and attractive.
Group uses class time effectively. All are involved.
Group asks good questions of each other on team.
Group asks good questions of other exhibiting groups.

"B" grade:

All of the above are fulfilled almost thoroughly.

"C" grade:

One of the above is not fulfilled.
Either one or two people in the group are lost; or
diagrams/maps are incomplete, messy, inaccurate;
or group does not stay on task in class.

"D" grade:

You don't know what you are doing.
You have omitted a required map or diagram.
You do not teach the class very much.
You dress funny.

Group 1. What are the forces that shaped the land?

Your group should construct detailed and instructive works for the following. You are also expected to understand the meaning of these diagrams. What do they show?

1. Schematic diagrams showing the cross sections of energy forces below the surface of the land (volcanoes, trenches, etc.)
2. Maps showing the land movement of the plates of the earth throughout earth history, focusing especially on the movement concerning Latin America.

Among the things you will need to find out:

1. Why are there mountains and volcanoes where they are in Latin America? (Why is this country full of highlands while Africa was mostly a land of plateaus?)
2. What are the natural hazards of the land? Why is there a persistent threat of earthquakes? What has to be done or has been done to accommodate this natural hazard?
3. How did the bridge between the Americas form? How has the movement of the earth's plates effected migration in Latin America (plant and animal)?

4. How does the earth produce the energy needed for all this colossal movement?

Group 2. What are the resources of the land?

Your group must find answers to and understand the following:

1. What are the animal, vegetable, mineral resources of the land? (Any oil?) How rich is the land for farming? Is there enough water? What food do they rely on and in what parts of Latin America do they use certain kinds of food?
2. Why do they have the kind of vegetation they have? Or, why are (or are not) the climate zones dry like Africa, which falls roughly on the same equatorial line and latitudinal lines as Latin America?
3. How do the seasons differ from those of North America?
4. How do Latin American birds and plants differ from those of Africa? Why do they have the adaptations, the differences they do? What special purpose do the unique plants of Latin America serve?

Your group must construct, and be able to discuss the meaning of, maps or diagrams showing the following:

1. The vegetation/climate zones in Latin America.
2. The atmospheric currents and important ocean currents which influence climatic zones in Latin America. You may need to include average rainfall statistics.
3. The hydrologic cycle.
4. The important resources and where they are found.

Group 3. Where do the people live on the land?

Your group is responsible for finding out:

1. Where did the first societies (and first migrant people) live in Latin America? Why there? How could people have migrated to Latin America? Could people have come from Africa?
2. Where was the population living around 1800-1850? What groups were living where? Why there? (Consider especially the groups/races of people throughout the West Indies and all of Latin America.)
3. Where do people live today? Why? What are the different groups/races living in Latin America today? Where do different language groups/races live today? What effect does that have on the countries in Latin America?

Your group must be able to construct and fully explain the following:

1. A map of the populations for each of the three questions above. Question 3 may require more than one map if you think it is necessary, or a clear overlay.

Thanks to Chet Pielock and Loretta Brady at Adelphi Academy, Brooklyn, New York.

in a summer workshop on exhibitions for Essential School teachers. "To ask about validity is to ask if the task represents the *real thing* we want to assess. Does it really present the student's abilities, traits, capacity for long-term work? For example, the SAT is valid because it statistically correlates with later success in college. But does it really represent the things the student *can* be good at, or just one thing?"

Reliability is another question, says Wiggins. "Would the student get the same score if he took the test again and gave the same performance?" he asks. "Or would different people score it differently? Standardized tests are reliable by design, but we question their validity. Exhibitions, on the other hand, are valid—but not necessarily reliable. How do we protect students from capricious, biased judgments?"

A related question is whether there must be one standard for the success of a student performance. Should standards vary according to the performer's level of intelligence,

When students learn to assess themselves, they internalize the criteria and become better performers and critics.

age, sex, race, family circumstances, future plans? "The failure to think through this question has led to us having no standards at all," Wiggins argues. "You can walk into any high school in America and see two teachers grade the same level work in dramatically different ways."

But teachers who use exhibitions in the classroom speak in matter-of-fact terms about how they evaluate student performance. "You've got to decide what's being graded ahead of time, and be clear about it with the student," says Melinda Nickle. "We assess the way they work, the way they use their time, the way they speak and write, the

ideas they bring to the performance—things that cannot be evaluated by a typical pen and paper test." Moreover, Nickle says, students are usually working in groups as they prepare their performances, freeing her to circulate, ask questions, and ascertain weaknesses and strengths.

"I don't question the accuracy of our assessment," Nickle says. "It actually is a lot more valuable than the traditional test, where what you mostly find out is if the student can memorize well or if he studied the night before. In fact, many students new to our program are unsettled by how high our expectations are—it is hard to get by without getting actively engaged in the learning."

Nickle is one of many teachers who require students to participate in their own assessments. "When a student presents an exhibition in my class," she says, "they might start out mumbling or speaking too fast. I say, 'go slow, breathe deeply,' and remind them that they are practicing speaking skills. Pretty soon the other students are prompt-

Figure 4. A Final Performance in History and English

Your final exhibition to demonstrate mastery of the material of these two courses for the first semester will be divided into two parts. The first part is a research paper. The second part is the final examination. Together these constitute 25% of your grade for English and 20% of your grade for World History.

1. For the research assignment, write a five to seven page paper addressing the following: "What matters in history is not societies or events, but individuals." Discuss the validity of this concept of history by citing at least three specific examples from your studies this semester that support or refute the concept.

Your paper will be graded by both your history and your English teacher and a grade assigned by each, based on the standards in the "Written Exhibition Assessment Form," attached.

2. The final examination will be taken during the period scheduled for the English examination. It will consist of three parts:

a. An essay evaluating your research paper, both in content and mechanics. You will read the attached excerpt on subjectivity, objectivity, relativity, and balance in academic writing (pages 6-7 from Toby Fulwiler, *College*

Writing. Boston: Scott, Foresman, 1988). As you read it, think how the points he makes apply to his research paper. Then write an essay reflecting on how these ideas are illustrated by your paper. Specifically, you need first to prepare a topic outline, including a thesis statement, for the essay. Then summarize in your own words each of Fulwiler's main points, and cite at least one specific example from your research paper of each of these points. Where you identify subjectivity or use of judgment in your paper, discuss whether there was adequate evidence in your paper to support these subjective statements. Finally, discuss why you think you made these particular subjective statements. In other words, how did one of your personal values enter into the research and writing of the paper?

b. An essay relating your English course readings to the thesis of your research paper. This will also involve an analysis of a short, related work during the final.

3. During the period scheduled for your history final examination, you will meet with both your history and English teacher to discuss your course work for the semester, your research paper, and your final examination.

Thanks for this performance to John Bohannon, a history teacher at Vermont Academy who attended a CES exhibitions workshop. (Vermont Academy is not a member of CES.)

ing them, too." Math teacher Glynn Meggison at Fox Lane High School in Bedford, New York has begun to invite his classes to grapple with self-evaluations. "One kid broke it down into actual percentages: quality of group work, individual work, presentation," he says. Such activities are themselves a form of learning, teachers say; as students internalize the criteria, they become better performers and better critics of others' performances.

The question of skill levels points up both a practical problem of comparative grading and a fundamental issue in education: the tracking of students early on into ability groups that will classify them for years to come. No one quarrels with the reality that students present themselves at different levels; but Wiggins argues that conventional means of grading—the bell-shaped curve, norm-referenced standardized tests, and tracking—merely discourage students from reaching toward higher goals. Again, he contrasts the academic model with the world of sports or the arts, where expert players are always before students as models of the excellence they are striving for.

Instead of giving watered-down challenges at lower skill levels, the same tasks are presented to all students, just as they are in real-life situations.

At Central Park East Secondary School in New York, students in a class of mixed ability levels set goals with the teacher ahead of time to aim for either a "competent" or "advanced" level of classroom performances, and they are evaluated accordingly. Instead of presenting watered-down challenges for lower

Figure 5. The APU Assessment of Mathematics (Great Britain)

The following section comes from the assessor's manual in an oral mathematics test of 15-year-olds, involving the ideas of perimeter, area, and circumference:

1. Ask: "What is the perimeter of a rectangle?" [*Write student answer.*]
2. Present sheet with rectangle ABCD. Ask: "Could you show me the perimeter of this rectangle?" *If necessary, teach.*
3. Ask, "How would you measure the perimeter of the rectangle?" *If necessary, prompt for full procedure. If necessary, teach.*
- ...
10. "Estimate the length of the circumference of this circle."
11. "What would you do to check your estimate?" (*String is on the table.*) *If no response, prompt for string.*
- ...
13. "Is there any other method?" *If student does not suggest using $C=\pi d$, prompt with, "Would it help to measure the diameter of the circle?"*

The scoring system works as follows:

- 1 unaided success
- 2 success following one prompt from the teacher
- 3 success following a series of prompts
- 4 teaching by the tester; prompts unsuccessful
- 5 an unsuccessful response; tester did not prompt or teach
- 6 an unsuccessful response despite prompting and teaching
- 7 question not given
- 8 unaided success where student corrected an unsuccessful attempt without help

Successful responses are combined into two larger categories called "unaided success" and "unaided plus aided success," with percentages given for each.

From Mathematical Development, Secondary Survey Report #1, Assessment of Performance Unit (APU), Department of Education and Science, Great Britain (1980).

skill levels—"dumbing down" the tests, as Grant Wiggins calls it—the same tasks are presented to all students, just as they are in real-life situations, and extra help given where necessary. The student can thus *learn from tests*—and learn, as he will in real life, from the more sophisticated responses of those at more advanced levels.

With clear objectives for pupil achievement, teachers can be more confident that grades really give direct information about where the student stands at that moment and where she is going next. This approach has been probably most carefully worked out in Great Britain, where a comprehensive national curriculum and testing

system has been under development for some time. The system provides step-by-step instructions for teachers to assess students at varying levels of ability. These guidelines actually ask teachers to prompt students if they cannot answer a question on their own—and if necessary, to *teach* the appropriate skill then and there. (See Figure 5.) Answers are calibrated on a scale that divides them into two large categories, "unaided success" and "aided success," within which students are rated further according to the level of their performance.

It seems clear that the careful devising of a scoring rubric is a crucial step in beginning to

evaluate student performances. One characteristic of a good scorecard is that it honors a variety of aspects of the student's performance. Is the student's work *process* being evaluated, for example, or merely the *product*? Are enterprise, flair, and creativity given equal weight with perseverance and carefulness, and are those weighted equally with achievement? Does the fluent speaker have as good a chance to excel as the fluent writer, the creative artist as good a chance as the computer whiz? The best performances are authentic reflections of a student's development of thoughtful habits of mind; they honor and use that student's unique qualities rather than force them into

One characteristic of a good scorecard is that it honors a variety of aspects of the student's performance.

a predetermined mold. (Richard Stiggins of the Northwest Regional Educational Laboratory has developed a short manual on how to design and develop performance assessments; it can be had for \$1 by writing NCME, Teaching Aids ITEMS Module #1, 1230 17th St., NW, Washington, DC 20036.)

Can You Fail an Exhibition?

In many ways, performances and exhibitions are set up so that a student cannot fail. If a classroom performance is inadequate, it serves not as a final judgment but as a revealing indicator of where the student needs extra attention in developing skills before the next performance. Evaluations are conducted on a continuum. As for culminating or graduation exhibitions, one simply does not attempt them until one is ready to do well.

In an authentic performance,

Grant Wiggins says, the student knows the nature of the challenge ahead of time, as with an athletic or artistic event. "The recital, debate, play, or game is the heart of the matter," he asserts. "All coaches happily teach to it." Because they can be practiced for, performances take on a teaching function at least as important as their evaluative function. And because they represent developing skills, a student's progress is emphasized rather than a scorecard of his errors.

"People do fail," Ted Sizer says. "People do persist in thinking that two and two makes five, or in writing graceless prose. At the same time, the same person who can't write may be able to draw—may be able to explain something of importance in a different way. The trick in good schooling is not just to meet minimum standards but to find out where a person is strong and make sure the person succeeds there. You need to see the best of a student on a regular basis; otherwise he loses self confidence."

Mike Goldman, who taught humanities at Central Park East and now works at CES, puts it even more strongly. "How can you test whether a student feels good about learning?" he asks. "That is the job of schools—and if a student doesn't have it, the *school* has failed, not the student." Grant Wiggins also questions what he calls the "gate keeping" function of assessment, which requires the student's performance to be reduced to one score and ranked. "Why must the transcript reduce performance to a letter grade that tells neither what a student can do nor what the strengths and weaknesses are that went into the grade?" he says. Moreover, some important things, like cooperation, may be best assessed indirectly, through other kinds of performances. "These things are complicated to assess—they involve values, attitudes, and the like," Wiggins says. "Do we want to be in the business of giving someone a grade on how much he

loves learning? Are we going to fall into the trap of the standardized test, that you can isolate every little thing and test for it?"

Recording Achievement

Developing a precise scoring rubric, of course, is not the only way of keeping records of student achievement. Alternatives include the keeping of anecdotal records, or assembling portfolios of a student's best work. At School Without Walls, a public alternative high school in Rochester, New York where performance is a fundamental aspect of instruction, teachers and community-based supervisors keep written records of student projects that become a permanent part of their records. Such anecdotal records can provide a richness of detailed information that cannot be achieved with any rating scale. In fact, at Central Park East principal Deborah Meier is experimenting with purely narrative records as an alternative to report cards that formerly combined comments with grade classifications like "Satisfactory" (plus or minus), "Distinguished," and "Unsatisfactory." "Otherwise parents and students zoom in on the grades, and ignore the rest of the report," she says.

In evaluations of student writing proficiency, portfolios have been

Send us yours . . .

CES is assembling a collection of exhibition and performance assignments to make available at Coalition workshops. Send material to Jill Davidson, Box 1938, Brown University, Providence, RI 02912. Please include your name, address, and telephone number for factual verification.

Portfolios don't have to be just writing; some students keep an ongoing videotape of their progress in various areas.

used increasingly by a number of states over the last decade. But student portfolios can include much more than writing samples; at School Without Walls a student's final transcript is a collection of teacher evaluations, critiques by community supervisors, state competency tests, and student writings. And at Thayer High School in Winchester, New Hampshire, principal Dennis Littky has students keep an ongoing videotape of their own progress in various areas from the time they enter school to their graduation.

How can teachers keep assessments from being capricious and biased? Grant Wiggins suggests that schools develop an oversight process such as a teacher committee on testing standards; in the process, he notes, useful dialogue and teacher training could be taking place. Where more standardized but teacher-given tests are being developed, as in Great Britain, a "group moderation" process allows collective scrutiny and review of any discrepancies that show up between different schools, or between one school's results and those on a national level.

The Measure of Competency

Still, the issue of measuring competency is real, and any school incorporating performances into its philosophy will need to face it. The answer seems twofold. First, at most Essential schools students who learn through regular performances actually do better on competency tests than they did in the old days. Figures from schools like

Walbrook High School in Baltimore bear this out (see *HORACE*, Volume 6, No. 1), and so does research into the effects of raised expectations on student motivation and performance. In a good exhibition, depth of coverage is more important than breadth; and to attain that depth, students must not only know their evidence, but know how to use it critically.

Second, some schools have been able to substitute their own demonstrations of mastery for state competency tests in various disciplines. States like Connecticut, Vermont, and California, and districts like Cincinnati, Pittsburgh, and Shoreham-Wading River, New York, are taking steps to incorporate performance-based assessment into not only writing evaluations but those for math, science, social studies, and other subjects. The trend is slowly making its way upward: at the federally supported National Assessment of Educational Progress, researchers are beginning to incorporate portfolios and other performance-based information into a broader base of information they have collected on student skills over the last two decades.

At the policy level, experts say, it will take a lot more than this to change an entrenched system that values easily measured skills over inquiry and thoughtfulness. Until policy makers issue clear directives for more performance-based assessment, schools will have to push for recognition of new ways on a case-by-case basis. But some signs indicate that organized teacher pressure is mounting for change at the policy level. English teachers were directly responsible for states' acceptance of portfolio-based writing assessment over the last decade. And recently the National Council of Teachers of Mathematics has introduced sweeping new demands in its criteria for curricula, focusing on mathematical challenges that push students to apply known facts in new ways.

Exit Exhibitions

Ultimately, Ted Sizer believes, the question of performance-based assessment must be addressed at an even broader and deeper level. If a school believes its chief task is to help students master thoughtful habits of mind, then the demonstration of that mastery—not the accumulation of credits, or the passing of time—must be the sole criterion by which students qualify for graduation. Many schools are moving toward this kind of goal when they require, for example, a senior thesis that integrates research in several disciplines. But a handful of Essential schools—among them Walden III in Racine, Wisconsin, Central Park East Secondary School, and School Without Walls—have gone much further, shaping their entire curricula so that the graduation exhibition, or "exit exhibition," is the focus of students' last years in school.

In the end, students who present an exit exhibition are asked to show they know much more than those in conventional schools.

What exactly is an exit exhibition? Ideally, it is a demonstration by the student in front of a review committee, at which he or she shows off the essential skills learned in the high school years. If the school's standards are met in the opinion of the committee, the student will receive the diploma. First, though, the student must stand up to the kind of probing questioning that we usually associate with the defense of a doctoral dissertation. It is not enough to show simple recall of memorized facts; the review board is looking for an ability to use