

To cite this article: Pat Hutchings & Ted Marchese (1990) Watching Assessment: Questions, Stories, Prospects, Change: The Magazine of Higher Learning, 22:5, 12-38, DOI: 10.1080/00091383.1990.9937653  
To link to this article: <http://dx.doi.org/10.1080/00091383.1990.9937653>

---

# WATCHING ASSESSMENT

---

# QUESTIONS, STORIES, PROSPECTS

BY PAT HUTCHINGS AND TED MARCHESE

**F**or the past four years, the two of us have been watching assessment—as observers and critics, as collectors of stories and documents, and through campus visits, fifty in all. Here, we set forth our sense of what this complex, often puzzling movement is about; how it plays out on campuses, including four we describe in detail; and what its prospects may be for lasting effect on undergraduate education.

Our view of assessment is mixed. As a phenomenon, it is at once powerful, scary in the wrong hands, increasingly a matter of law, and home to the day's most provocative discussions of teaching and learning. As a movement, it tilts at the deepest structures and habits of academic culture; its practice on campus is marked by tricky beginnings and important accomplishments. The one sure thing is that assessment warrants close attention. Here's how we see it.

---

*PAT HUTCHINGS chaired the English department at Alverno College for nine years before coming to Washington, D.C. in 1987 as first director of the AAHE Assessment Forum. TED MARCHESE taught and held administrative posts at Barat College before coming to AAHE as vice president in 1982; he has been an executive editor of Change since 1984. Hutchings and Marchese have collaborated on assessment work for four years; the Forum itself and the field visits behind this article were supported in part by a grant from the Fund for the Improvement of Postsecondary Education.*

## I. Hard Questions

**G**et yourself into the middle of talk about assessment and you're likely to hear a lot about methods, mandates, and measurement. You'll hear from people who come at assessment as program evaluation, quality control, or psychological testing; as a tool for management, a classroom strategy, institutional research, or public relations.

But assessment is best understood as *a set of questions*—questions that are not, in fact, entirely new, but that now come at us with greater insistence. At bottom they're questions about *student learning*:

- What is the college's contribution to student learning? How and what do we know of that contribution?
- Do our graduates know and can they do what our degrees imply? How do we ensure that?
- What do the courses and instruction we provide add up to for students? Are they learning what we're teaching?
- What knowledge and abilities do we intend that students acquire? Do they have opportunities to do so? Are they successful? At what level? Is that level good enough?
- How can the quantity and quality of student learning be improved? What

combination of college and student effort would it take to achieve higher levels of performance?

These are good, important questions; we need to be able to answer them. They're also hard questions, not the least because they bring to the surface fundamental issues of institutional purpose.

They are questions, we've found, that faculty understand—questions good teachers ask about their own students. Is this student getting it? How do I know? What constitutes acceptable work? What can I do to help my students succeed? "What I understand by assessment," a community college faculty member wrote to us, "is asking, 'Are my students learning what I think I'm teaching?'" Faculty ask assessment's questions because they need answers to do their daily work with students. Assessment, for faculty, means a habit of looking at teaching in terms of its effects: learning.

Individual faculty are not alone in asking such questions. Departments and programs, even entire institutions, are learning to ask about their cumulative impact on student learning—about "what it adds up to." What knowledge and abilities should (and do) our accounting majors have when they graduate? In what ways do all of the teaching and course-taking in biology come together—or not—for students? What are the outcomes of our general educa-

tion program? Do the recipients of our degrees in fact possess the traits of intellect and character the college promises in its catalog?

Not surprisingly, assessment's questions ring loud bells for higher education's outside constituencies. The public at large retains a faith that higher education is a good thing, something it wants for its children. But there's a sense, too, that things aren't quite right on campuses, that a great deal of money is being spent to uncertain effect.

Policymakers have questions particularly about undergraduate teaching and learning, which they see as neglected. At a recent meeting of state leaders in Santa Fe, New Mexico, Joyce Holmberg, an Illinois legislator, noted, "I get constituent complaints all the time that the best professors are off doing research, not teaching at the lower division. The quality of teaching in our universities is a real problem. Tell me how to fix it." To Holmberg and others—in all but about 10 states now—assessment looks like part of the answer; its questions are their questions.

It need hardly be said that legislators' rendition of those questions—often embodied in mandates—differs in tone from the versions asked by campus administrators, whose versions differ yet from those of faculty members. There have been hard feelings, tension, and lots of frustration on all sides. Getting the various levels of questions to connect rather than collide has proven a major stumbling block.

But with the noise turned down it's possible to hear a common note as well: that when it comes to higher education, we *all* need to ask not just about the funding, facilities, credentials, and curricula but about *results*. Assessment's questions raise that issue; they make *student learning* the result that counts in gauging institutional performance.

## II. A Tradition of Questions Lost...

Assessment's questions are far from new; indeed, in asking about the cumulative effects of college and a given student's readiness for a degree, assessment points back to earlier ways of thinking about educational quality.

In the 19th-century American college

## Assessment Hits Home

**Eighty-two percent of all colleges now report "assessment activities underway," according to 1990 Campus Trends data from the American Council on Education. That's up from 67 percent in 1989 and 55 percent in '88.**

**In prior years, activity has been higher in public institutions than in the independents (by 79 to 56 percent in '89); now the two sectors run neck and neck, at 79 percent each, with public two-years at 87 percent. Some 42 percent of reporting institutions say their state requires assessment; over half describe their activity as "part of a self-study for a regional accrediting agency."**

**Attitudes toward assessment continue to be mixed. Half of all respon-**

**dents believe "student assessment will significantly improve undergraduate education." At the same time, 73 percent express "fears about misuse of effectiveness measures by external agencies."**

**As to methods and approaches, 66 percent report reliance on their own, locally designed instruments. Seven in ten report concern that nationally normed standardized tests risk "distorting the educational process." One in three is now experimenting with portfolios, a method one heard little about two years ago. (Source: *Campus Trends*, edited by Elaine El-Khawas, vice president for research and policy analysis, American Council on Education.)**

it was assumed that the mere taking of courses was insufficient warrant for a degree. A candidate for the bachelor's, therefore, faced a final hurdle of the senior declamation, a requirement that he or she publicly demonstrate a full possession of knowledge and high skills of intellect and speech. Often these examinations were conducted orally, by and before outsiders, sometimes clergy, at Amherst by "Literary Gentlemen of Good Standing." Their point, in modern terms, was quality assurance before relevant communities. Underlying the practice was an assumption, carried over from English universities, that instruction and evaluation ought properly to be distinguished and be done by different parties. The exercise, too, was thought good for students.

In the closing decades of last century and the early years of this one, the American college became today's university. Along came an elective system and a great expansion of subjects for study, huge increases in student numbers, the ascendancy of research and graduate study, faculty authority and autonomy. The nature of faculty work and role also changed, from the English collegiate model to a continental-university model in which "good teaching" meant knowing one's subject deeply and speaking clearly about it, with learning seen as the *student's* responsibility. This narrowed conception of faculty role, combined with an explosion of curricula and student numbers, made the practice of comprehensive senior exams—the asking of questions about the cumulative effects of college course-taking on student learning—increasingly difficult to sustain.

This did raise a new problem: on what basis to award the degree. In the 1880s, a bit of administrative invention filled the breach: instruction and evaluation (grading) were combined within individual courses, each course was assigned a credit-hour value, and when the credits (C average or better) totaled 120 the degree would be issued. Not without protest did this all happen. As late as 1911 Harvard president A.L. Lowell addressed the Association of American Universities on the topic, "Disadvantages of the Current American Practice of Conferring Degrees (With the Exception of the Ph.D.) on

the Accumulation of Credits in Individual Courses, Rather Than As The Result of Comprehensive Examinations upon Broad Subjects."

The 1930s were a period of intense questioning and innovation in American higher education. For reformers like Abraham Flexner and Frank Aydelotte, restoration of the comprehensive, senior-level exam became a key plank in the platform for change; through such exams, reason and integrity might return to curricula and competence attach to the baccalaureate. A 1935 survey counted 242 colleges that required some sort of "senior comp," up from 71 in 1925. Except at smaller liberal arts colleges, however, most such exams were restricted to the major, where scale and common studies made the device more workable; questions about larger cumulative learning tended to be left aside.

It has been many decades since the comprehensive examining of seniors played a significant role in the award of degrees. In losing that practice, we lost as well a tradition of asking questions about our graduates' competence and about the cumulative effects of our teaching and curricula.

### III. ...and Questions Found

Now, thanks to assessment, these questions are with us once again. Why now? One answer is that these things go in cycles. Carnegie Mellon historian Daniel Resnick has chronicled large swings over the decades between attention to access and system expansion and then, predictably, to consolidation and questions of quality. The early years of this century were ones of great increases in schooling at all levels; the '30s, when times were harder, evoked many of the questions we hear now. Education at all levels grew enormously through the baby-boom decades; now again, with a pinch on the economic side, come questions about system performance and student learning.

Whatever the cycles, powerful ideas also lie behind today's call to assess. Ten years ago UCLA's Alexander Astin recast debate with an argument that traditional ways of thinking about quality in higher education—quality as a function of *resources* (high student

*"There is a conflict of interest in the way in which American colleges and universities certify instruction...*

*Faculty members not only teach but in effect guarantee, first, that their teaching meets established standards in both content and quality and, second, that students have learned what faculty have taught. There is no external mechanism to verify the integrity of the baccalaureate degree."*

—Joseph P. O'Neill,  
Scholar in Residence,  
Carnegie Foundation for the  
Advancement of Teaching, 1983.

*"Either faculty must become a great deal more sophisticated and rigorous in their system of evaluation, or evaluation by units external to the classroom will increase."*

—John Losak,  
Dean, Miami-Dade  
Community College, 1987.

SATs, faculty Ph.D.s, endowment, library holdings) or even of right *processes* (such as rich curricula and good advisement)—told too little, misled even; that the real measure of quality was found in a college's results, its contribution to student learning, the "value added" from the experiences it provided. Outcomes mattered in this view, as did attainments over time, and *evidence* of the two—an agenda for assessment—was advanced as a necessity.

By the mid-'80s, Astin's views had taken hold in an undergraduate reform movement growing within the academy, spearheaded by two influential reports. In late 1984, a National Institute of Education study panel (on which Astin sat) issued "Involvement in Learning," which argued that to

*"In any normal collegiate program each student will have to face thirty or more courses, each with some form of final examination or paper. In other words, he or she is asked to rise to a bewildering variety of challenges, and to concentrate both knowledge and skill to meet them... Freshmen twitch about examinations and final papers; seniors take them in stride. That stride is a clear 'value added' by the four years of college."*

—Father Timothy S. Healy,  
writing as president of  
Georgetown University, 1987.

*"Effective and even inspired instruction in isolated courses does not constitute good education. A faculty member must see his/her role as one of a team composed of an entire faculty dedicated to developing a series of interrelated and coherent educational experiences which stimulate and assist the student to become a self-educating individual."*

—Paul Dressel and Lewis Mayhew, 1954.

strengthen learning one needed to a) involve students in their studies, b) set high expectations, and c) assess and provide feedback. In early 1985, the Association of American Colleges' "Integrity in the College Curriculum" also made this learning-assessment link, calling it "scandalous" that colleges failed to assess the impacts of their teaching. Behind both reports lies a view that quality is indeed a function of student learning.

A subtle but important development over the next two years was the way this view of quality gained currency within the academy. It came to permeate convention speeches, journal articles, faculty debates, and, late in 1986, a book by the president of Harvard, Derek Bok's *Higher Learning*. In varying tones, William Bennett, Ernest Boyer, and a host of others gave further voice to "the need to assess."

It was in part from educators themselves, then, that state policymakers took their interest in assessment. Spurred by concerns about "economic competitiveness" and "workforce capability," the education-minded governors of the mid-'80s began asking now-familiar questions about student learning and undergraduate performance. Missouri Governor John Ashcroft, in his capacity as chair of the National Governors' Association Task Force on College Quality, put it bluntly in 1986:

The public has a right to know and understand the quality of undergraduate education that young people receive from publicly funded colleges. . . . They have a right to know that their resources are being wisely invested and committed. . . . We need not just more money for education, we need more education for the money.

Many governors (and legislators) turned to assessment because it was a "solution" they recognized from the K-12 level. They had spent the early years of the '80s asking hard questions about performance in the schools; a policy response in all 50 states was to impose statewide standardized tests at various grade levels. States call this testing "assessment," and they believe it works. When attentions turned at mid-decade to higher education, some legislators thought they spotted a familiar problem (inadequate student and institutional performance) and looked to

a familiar solution, "assessment" as testing.

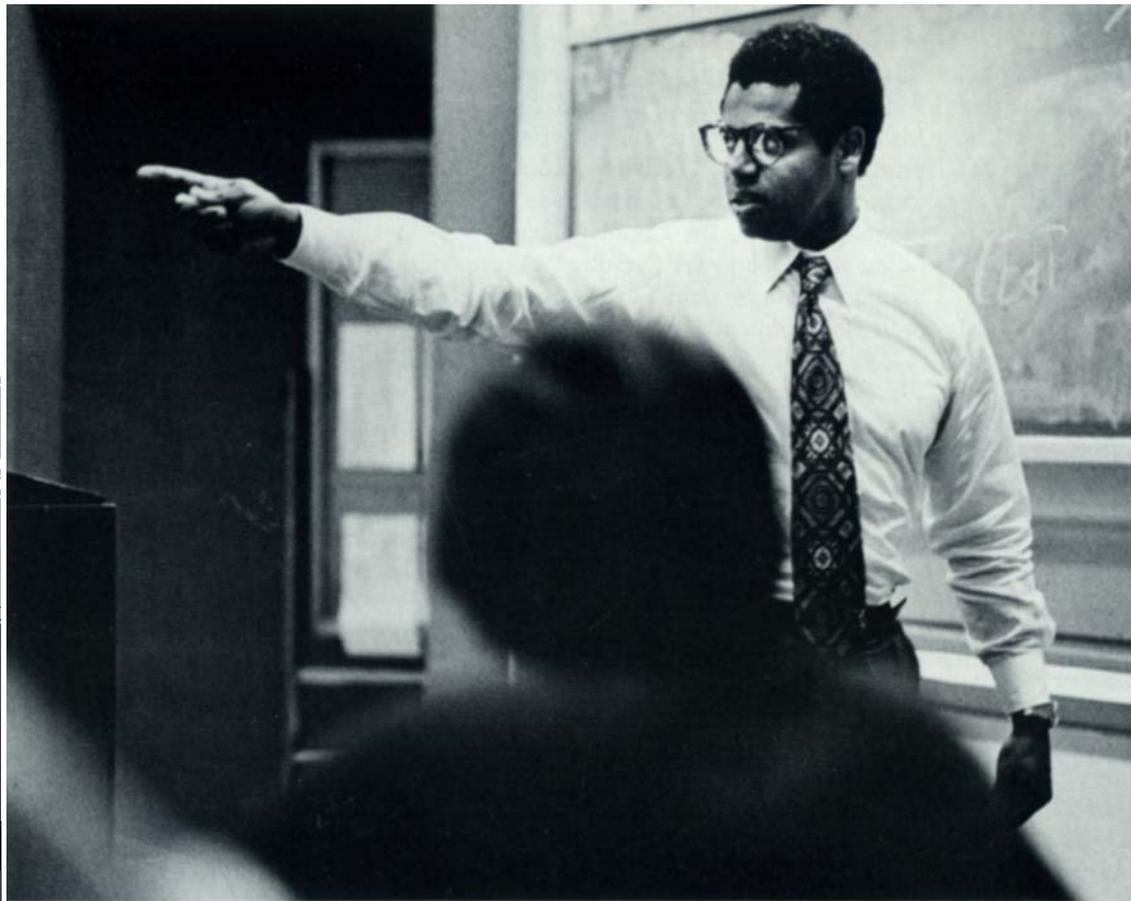
Adding to their interest was the visibility of several successful assessment programs in higher education. Alverno College and Northeast Missouri State University, both of which had used assessment (albeit in very different ways) to improve student learning, were glowingly profiled in Ashcroft's NGA report, suggestively titled "Time for Results."

A combination of ideas, then—that quality was a function of learning, that learning should be demonstrated, and that assessment could be a lever for improvement—undergirded state interest in postsecondary assessment.

That interest has now moved assessment across the threshold from "another interesting idea" to a "condition for doing business." While at mid-decade just three or four states took an active role in promoting assessment, some 40 states are now taking steps to require or promote assessment. External pressure comes too from a second source: accrediting agencies face U.S. Department of Education (1988) and Council on Postsecondary Accreditation (1987) rules that stipulate, as a condition for their own approval, that they must require information about learning outcomes from the institutions and programs they review.

Judging by recent surveys of governmental activity on the issue (it is higher than ever), public interest in assessment will endure. That's likely, too, because the civic and economic imperatives behind the demands for educational performance are worldwide and can only grow. If you talk with legislators and lay trustees, you know they *like* assessment's questions, and want them answered; they see them as questions about quality improvement, not unlike those they pose to other state agencies and services or back in their own businesses. Indeed, they seem surprised that colleges haven't *routinely* sought evidence about impacts and effectiveness, and see the absence of assessment as self-serving protectiveness.

As late as 1987 one could still hear assessment called a fad; but no more. To many observers—certainly to us—assessment's questions seem sure to be with us in a big way for a long time.



*The University of Virginia faces a state mandate for assessment "appropriate to its mission" and reflecting a "diversity of program goals."*

#### IV. A Tale of Two Campuses

What it means for a campus to come up against these questions, to find a place for them within its own agendas, traditions, and daily life, can best be understood through the scores of campus stories we've watched unfold. To convey some of the textures and themes of these stories, we recount here visits to two flagship state universities where assessment has recently arrived.

##### *The University of Virginia*

A chamber of commerce brochure calls the Charlottesville campus of the University of Virginia "one of America's top 10 travel destinations." It was easy to appreciate that designation on our two-day visit late last fall: clipped lawns splashed with leaves from big old trees, firewood at the door of senior rooms on the Lawn, Thomas Jefferson's Rotunda in sparkling shape after UVA's recent hosting of the "Education Summit" attended by President Bush. Ted Fiske's *Guide to Colleges-1989* gives UVA fourteen stars (only Stanford and Brown have as many);

13,386 students, 7,900 of them from out of state, stood in line for a place in the previous fall's freshman class of 2,626. Ninety-seven percent of those entrants will return for the sophomore year; 90 percent complete their degree.

With so many signs of wellness at hand, UVA administrators and faculty

were understandably perplexed by the state's insistence that they do assessment. In different tones and words, a refrain we heard on campus went like this: "We know Richmond wants to improve undergraduate quality across the state. . . . We're willing to go along and do assessment, but, *really*,

### State Trends

As late as 1985, the universe of states active in promoting assessment numbered just three or four; by 1987, that number had grown to a dozen; today it is close to forty. Earlier fears (circa 1985-87) that states would roll out mandatory statewide tests have not been borne out. Instead, once institutions caught on and got their arguments in line ("there is no collegiate equivalent to a universal math test for ninth graders"), some two-thirds of the states chose to follow the more permissive path charted by Virginia and Colorado: each public institution is to "do assessment" in ways of its own choosing, consistent with its particular mission and clientele, with required reports focused largely on evidence that it has put

findings to use.

Reflecting the current mood, three-fourths of the states polled last winter by AAHE and the Education Commission of the States averred that "institutional improvement," not "accountability," is the main point of their assessment initiative. How long will that last? One state-board chief put it this way: "It all depends on how the institutions respond . . . if they don't take the process seriously, there will be an enormous pressure to centralize."

For information on the poll and related papers about state policy on assessment, write: Joni Finney, Education Commission of the States, 707 17th Street, Suite 2700, Denver, CO 80202-3427.

what could the state have in mind *here?*”

That refrain came with a certain poignancy, an urgency even. Since that summer (1989), the university and the State Council of Higher Education for Virginia (SCHEV, as it is called) had, after some progress, come to loggerheads over assessment. Here was a state board whose assessment guidelines had been held up as a model across the country, and a much-admired campus. What was the story?

Virginia's assessment initiative traces to a legislative resolution passed in 1985 directing SCHEV “to investigate means by which student achievement may be measured to assure the citizens of Virginia of continuing high quality of higher education in the Commonwealth.” In response, a statewide committee devised “Guidelines for Student Assessment,” ten in number, which SCHEV approved April 3, 1987. These abjured a statewide test (the fear, then) in favor of individual plans “appropriate to the mission of each institution” and that reflected “diversity of program goals.” The guidelines left wide latitude for institutions to devise their own approach; urged use of existing data and sensitivity to assessment's effects on students; asked for attention to general education, remediation, the major, and alumni follow-up; made clear that the object was not to compare colleges “but to improve student

learning and performance”; and required evidence that the effort was “useful” and linked to “programs to address identified areas of weakness.” Each institution was required to report yearly to SCHEV, first on its plan, then on its progress, and have these approved; by 1989, the Council's expectation would be to see “results,” these “concrete, more than anecdotal, and presented in quantified form.”

UVA's first plan was submitted June 29, 1987, and approved by SCHEV two weeks later. It called for the use of interviews to trace over time aspects of the “university experience” and student progress in general education; for studies of outcomes in the major to be piloted in a few departments, anticipating assessment's eventual use in program review; for assessment to be built into a plan for “exceptional” students; for alumni follow-up—in fall of 1991; and it promised that the provost would propose to the faculty a new requirement for assessment-oriented senior projects for all students.

That document was also a call for time, care, and caution. “The University of Virginia's reputation as an especially attractive place for undergraduates is not based solely upon architecture, atmospherics, and celebrations,” provost Paul Gross reminded SCHEV in his transmittal letter. “It reflects a national recognition that this place cares for its undergraduates and specif-

ically for their minds. . . . The interesting new [assessment] ventures we propose . . . are the set we believe can work here, with minimum disruption of existing systems and staff, and with maximum yield of interpretable information.”

Making this all happen within the parameters intended by SCHEV became the challenge. A committee was formed, but committees need time to find their direction in new terrain. Gross's attempts to interest department chairs in a required senior project met with resistance, mostly on grounds of cost; determining how to do the longitudinal study turned out to be difficult. UVA's request of SCHEV for \$35,000 in supplemental funding for pilot work was answered by a grant of \$3,300.

As difficulties mounted, momentum flagged and skepticism grew. On April 9, 1988, the chair of the Assessment Steering Committee, sociologist Steven Nock, voiced his own doubts in a paper for a faculty colloquium. It reproves SCHEV for mandating assessment without specifying *why* it was doing so, for failing to point to “some problem” that would help orient institutional response. “In what ways are we failing?” it challenges. “Indeed,” Nock adds, “SCHEV has never defined assessment.”

Our belief is that SCHEV assumes that the outcomes of education can be specified and measured. This, we suspect, is their definition of assessment: *the specification of desired outcomes and the measurement of students to determine whether such outcomes have, in fact, been imparted.* If so, this creates yet another dilemma for us for we have been unable to locate a literature which advises whether or not it is possible to define the desired outcomes of education and, if so, how such outcomes are to be measured.

The Assessment Steering Committee rejects the idea that outcomes of education can be agreed upon by educators—at UVA or anywhere. . . .

It was hard for us to reconstruct, after the fact, the degree to which Nock's attitude was shared and how it affected assessment activities. UVA *did* put together the required report to SCHEV on activities in 1987–88, and that report was accepted. But back came warnings. The letter from Margaret Miller, SCHEV's number-two officer and its

## Harvard University

In 1986, president Derek Bok told *Change* readers that every college should “study the learning process and assess the effects of its programs.” That year he asked Professor Richard Light to convene what came to be called the “Harvard Assessment Seminars”—monthly dinners followed by small-group work sessions—these “of, by, and for faculty,” eventually 100 people in all, who choose what to assess, then become a constituency for findings.

When Light brought together the project's first findings last spring (“Explorations with Students and Faculty about Teaching, Learning, and Student Life”), the report garnered wide media attention and 6,000 requests for copies, many

from faculty. The latter seemed a function of the report's concreteness in responding to teachers' questions: *Why do women drop out of science? What are the characteristics of highly admired courses? How can small groups be used to enhance learning?*

Harvard's approach pays little attention to outcomes or external reporting; its focus has been on “real questions that faculty want answers to” and on “locally useful information,” according to Light, who expects from it “small but steady increments of improvement.” The approach—“faculty conversation and inquiries into student learning”—fits Harvard's culture, has brought results, and is a suggestive model for sister institutions.

UVA PHOTO BY GARY ALTER



*UVA assessment plans include interviews to track students' experience of the university and their progress in general education.*

*"It is easy, in focusing on the internal-external tensions, to miss important points of agreement. Both institutions and external groups would like to have more and better assessment information. No one is sure how or where to get it, how to use it, or even whether or not a reasonably good system can be put in place."*

—William W. Turnbull,  
writing in *Change*, 1985.

head of academic programs, is pointed in asking about stretched-out timetables, scaled-back plans, and claims of insufficient funds; it raises detailed questions about research design. Several more letters—a growing coolness and impatience barely concealed—passed back and forth between Charlottesville and Richmond before all parties agreed to a September 29th face-to-face meeting on campus.

At that meeting, we're told, no party backed down, but peace pipes were smoked; Gross promised the university's best efforts to move ahead, and Miller pledged her help along the way. To speed implementation, UVa brought aboard Ned Moomaw, a political scientist and veteran administrator, as coordinator of student assessment.

Through '88-89, renewed efforts went into the launching of a "Longitudinal Study of Undergraduate Education," adapted from the work of Herant Katchadourian at Stanford, aimed at following in detail the experiences of students in UVa's class of '92. A 15-page questionnaire was developed that fall, administered to a fourth of the class (482 respondents) in March, after which about a third (155) of those freshmen were interviewed for an hour each by faculty members over the course of April and May.

We found the results fascinating and rich in implications. Among other things, questionnaire findings indicated very high student satisfaction with their overall experience at the university (approval ratings of 81 percent),

but trailed downward for satisfaction with teaching (65 percent), TAs (53 percent), contact with faculty (40 percent), and advisement (28 percent). In the interviews, freshman recollections of academic life center almost entirely on the mechanics of getting through classes; when asked, a few spoke of a memorable intellectual experience.

As is always the case with data, the meaning of these numbers—even whether they're good news or bad—is by no means self-evident; careful discussion and "meaning making" is needed. At the time of our visit, we found little evidence of such discussion, suggesting to us assessment's lack of a constituency on campus.

Meanwhile, Moomaw and the assessment committee found five departments interested in undertaking a trial assessment of their major. The resulting reports—from biology, environmental sciences, history, religious studies, and romance languages—could be faulted—as they were by SCHEV—as less than state of the art. Even so, we were encouraged by what we heard from the department chairs we interviewed, including stories about assessment's power to raise levels of introspection. "These are questions we should have been asking all along," James Childress, chair of religious studies, told us.

As for other activities: The alumni follow-up study remained for the future (now 1992-93). The required look at "developmental students" was addressed by looking at a subset of such students (an *n* of 17) within the longitudinal study. As for general education, the university argued that it too would be scrutinized via the longitudinal study. In July 1989, the university submitted its required annual report to Richmond—37 pages of assessment developments.

Alas, to the high dismay of parties in Charlottesville, the report—in effect—bounced. Back in no time came a 9-page letter from SCHEV's Margaret Miller. Reporting results of a review by the SCHEV staff and a national consultant, Miller expressed "disappointment with the general progress" of the university, which to the review team suggested "a lack of commitment to the overall . . . process and an absence

of effective leadership. . . ." The design and analysis of the longitudinal study were praised, but the team noted that it did not address "what students are actually learning." UVa was criticized for overly optimistic interpretations of the data and failure to address the problems it uncovered. As for general education, "the university has not made progress," its timetable is "too leisurely"; of the five departments, "none has an assessment program that is adequate." Efforts to know about remedial students were dismissed as too little, the timetable for alumni study too late. Most importantly, given the 1986 mandate's expectation for 1989, SCHEV found that UVa reported "too little in the way of assessment results."

In the genteel world of Old Dominion higher education, Miller's language sparked strong reaction. Provost Paul Gross, about to leave office at summer's end and finding little reason for restraint, fired back a two-page response: "I find myself at odds with your ideas . . . on what assessments are worth doing in a place like this, how much it is worth spending on the activity, and what relationship they can be expected to have to improved learning for *our* students." Gross took the additional step of sending copies of Miller's letter and his own to everybody (it seemed) in Charlottesville and beyond who had anything to do with assessment.

The atmosphere was still charged at the time of our visit in November. Given the turn of events, the former assessment committee had been disbanded, a new one formed; we found it somewhat dispirited and unsure of its role. "Some people here are mad at SCHEV and want to fight, others figure we might as well do whatever they want, regardless," Sam Kellams of the Ed school told us.

But to acting provost Hugh Kelly the issue was clear: "We've got to do what's right for us, but get things back on track with SCHEV." What was at stake, he knew, was the university's continued eligibility for state incentive funding—even, perhaps, its lesser sum for assessment (\$168,300 in 1988-89).

As we've stayed in touch with UVa developments, we've been greatly en-

couraged; assessment clearly turned a corner last winter in Charlottesville and has since moved ahead. Miller and Kelly have forged a new agreement to make assessment work. More flexible guidelines for assessment of the major have been fashioned and "most departments" are engaged with the matter. Fuller information on the experience of developmental students has been assembled; round two of the longitudinal study came off in style (94 percent of the sophomores participated); the alumni survey has been moved up a year; the General Education Committee has proposed a new program built around four skills and five perspectives, with objectives for each that anticipate assessment. In March, Kelly's office published and sent to the entire university community and to SCHEV a crisp, provocative report of findings from the longitudinal study, a step "that is already producing results" and that both Miller and Gordon Davies, director of SCHEV, responded to enthusiastically.

In a most interesting development—acting on its realization that there was no active conversation about undergraduate learning on campus and con-

sequently no "place" for assessment findings to go, be discussed, then acted upon—the assessment steering committee last spring adopted for itself the model of the Harvard Assessment Seminars, repositioning itself as a group that will foster "a broad-based and wide-ranging conversation about the quality of undergraduate education at the university." Subcommittees now handle the everyday work of the assessment program. Professor Richard Light, leader of the Harvard effort, was visited by Moomaw last spring; this fall, Light will be the new committee's inaugural speaker. How this approach will be received in Richmond remains to be seen, but SCHEV's initial reactions to UVa's most recent work have been positive.

Why this turn-around? Observers tell us that SCHEV, as it gained experience with assessment, became less centrally focused and more accepting of local circumstances, to the benefit of UVa. But the larger change doubtless occurred in Charlottesville.

Historian Alexander Sedgwick watched this story unfold during his deanship of arts and sciences. In his view, UVa's initial approach was to handle assess-

ment administratively, in effect to contain it, a strategy that collapsed when key parts of the program didn't fly—in Richmond or within the faculty. The result was the "confusion and hostility" of 1988 and 1989.

Kelly's approach when he took over last fall was to convince SCHEV that he was serious about assessment, then to "give assessment back to the faculty" via messages that it was a valuable way to energize teaching and curricula but that "the how" would be up to them. Release of the longitudinal-study report in March "made a difference right away" by showing faculty how valuable assessment information could be.

"In effect," Sedgwick says, "all of us—the administration, the faculty, SCHEV—finally got on the same wave length."

"Another perspective," Sedgwick told us, "is this. The previous provost [Gross] came into office with orders to emphasize research. A year ago, as Kelly came in, our Board of Visitors started to complain that we were neglecting undergraduate education. Suddenly, assessment changed from a problem to an answer."

## A Changing Scene

The now-annual conferences of the AAHE Assessment Forum offer a good gauge of assessment's progress. At the first such conference (co-sponsored by NIE and AAHE) in 1985, an overflow crowd of 700 struggled to understand just what this new thing called assessment was. The air was filled with "assessment is coming!" talk, dire predictions, arguments from different schools of thought, and brave plans to start this or that campus program.

By 1987 and the Second National Conference, talk had turned from "what" to "how." Another audience of 700, many of them administrators feeling the pinch of state timelines and requirements, came looking for "models." The Denver program featured a handful of such: the external examiner, new standardized instruments, computer-adaptive testing, basic-skills approaches, assessment-center methods. . . . Assessment was beginning to show signs of breadth, if not depth.

By 1988 in Chicago (the third gathering, with 1,000 on hand), greater thoughtfulness and experience were evident among practitioners. Many more examples and options were on show, and for the first time people seemed able to talk from actual practice about what worked, what didn't. On the latter front, frustrations with available instruments brought on a wave of test-bashing; the search for alternatives was on.

By the 1989 Atlanta gathering, those alternatives were clearly on the table—portfolios, interviews, focus groups, capstone courses with senior projects . . . and with them much greater confidence by campus teams that they could figure out a right thing to do. The 1,150-person audience (Peter Ewell described them as "happy amateurs") was a yeasty mix of faculty from a range of disciplines plus younger academic-affairs administrators; many of them seemed to care less about doing studies and collecting data

than they did about issues of undergraduate improvement and the routes to it opened by assessment's questions.

This past June's conference (the fifth) brought over 1,400 participants to Washington, D.C., many in large campus teams. Sessions were marked by candor about how assessment has unfolded (often not too smoothly) on named campuses. The roster also turned up a raft of "assessment coordinators," a position that barely existed three years ago. The dominant note of the event was an insistence that whatever is done in the name of assessment must be connected with teaching and learning.

Having attended four of the five conferences, E. Thomas Moran, vice president for academic affairs at SUNY-Plattsburgh, marvels at the progress they reveal: "Two years ago assessment practitioners worried that the movement had peaked and that we'd started to repeat ourselves," Moran noted at this year's event. "That obviously hasn't happened."

*Assessment initiatives at UConn raised valuable questions about the link between teaching, curricula, and goals.*



### *The University of Connecticut*

New questions were abroad on the Storrs campus of the University of Connecticut, too, when we visited early last winter. A driving rainstorm made it a wet slog from building to building as we interviewed assessment committee members, their critics, students, and top administrators about how assessment had unfolded there. Like Virginia, Connecticut has an assessment mandate, one that calls for each institution to come up with its own plan; unlike UVa, UConn's effort predates the mandate and has been faculty-driven.

That effort traces to 1984, when the university undertook a study of curricular options for general education. In 1986 that study led to the adoption of a model that built on the general education program already in place in the arts and sciences; the new model requires 36 hours of work (including non-Western studies) in six areas, the latter loosely defined and met by a broad spectrum of departmental courses. Starting in

1988, these new requirements applied not just to students in the College of Liberal Arts and Sciences (which enrolls 62 percent of the 16,300 undergraduates on UConn's six campuses) but to *all* undergraduates, including those in UConn's strong professional schools.

The Senate's passage of these new requirements in 1986 set in motion the assessment story we came to investigate. A moment after the decisive ballot, economist Peter Barth rose to question "how we'd know this new curriculum was working." On the spot a committee to evaluate the new curriculum was created, headed by geographer Judith Meyer.

The new committee's work was dogged at first by suspicions that its real agenda was to sandbag what had just been put in place. Good communication soon quieted those fears, leaving the committee to face a tougher problem: general education at Connecticut consists of some 100 different courses—actually less sprawling than at many similar places, but hardly a "program"

in the sense of having connected elements aimed toward intended outcomes. Indeed, other than one-sentence descriptors of the six areas of study, goals for the program had never been established, leaving little to go on for purposes of evaluation. Given their absence, the committee spent a full year articulating and refining goals, along the way involving dozens of faculty from relevant departments and colleges.

In May of 1987, after this year of work, Meyer attended a meeting called by the state's Department of Higher Education. On her return, she told the committee of the new word for what they were doing—"assessment"—which in fact the state would soon require; what's more, a federal agency called FIPSE had money for institutions that wanted to try the thing. Committee members led by associate professor of German Barbara Wright got a successful proposal together and soon the faculty group, with a \$148,000 grant in hand and Wright as project director, set to work on assessment at UConn.

Notably, the committee continued to see assessment, *nee* evaluation, as means, not end. Despite early misgivings about the committee's intent, it soon coalesced around a shared sense of need "to do a better job of general education for our students," in Wright's words. As the committee struggled with goals for the six areas, and still later wrote and administered batteries of home-developed tests, the expectation was not for an assessment breakthrough ("We're just a host of amateurs who set off to learn what we could," as Wright tells it); the hope was to end what Wright calls *tot-schweigen*, the wide indifference to general education, the tendency to ignore or "silence it to death."

A clear accomplishment of the work led by Meyer and Wright is that, over time, nearly 100 faculty members, many of them senior and from all departments, have been drawn into engagement with general education as an issue. As Lee Jacobus of the English department told us, he stays with the assessment group "because it's the only place on campus I can find an important conversation about what students are learning." At the end of our day-long visit, 150 faculty, seemingly appreciative, turned out for a two-hour report by committee members on work-to-date.

As at UVa, this work has not gone forward without its skeptics. Historian Anita Walker told us that "a department will offer eighteen sections of an

introductory course, each taught differently, so what can you assess?"

"I respect the judgment and academic freedom of the teachers of these courses, each of whom is different," said Mort Tenzer, a political scientist. Jacobus remarked, "The fact is, *courses* don't educate students, *people* do, faculty members, each in his or her own way. . . . When students take Milton or Irish Lit from me, they take Lee Jacobus. I think it's impossible to assess the college experience if you think of it as courses instead of as a range of experiences that vary by individual."

Also in doubt (and not without reason) was the possibility of separating out effects of general education from all other coursework. As Tenzer put it, "Gen ed is so intertwined with other instruction that there's no distinct thing there to evaluate." Further, in Walker's words, "How do you assure the quality of student effort in taking these tests?" "Make the tests fascinating!" Jacobus rejoined, not entirely negating feelings that something "uncontrollable" was afoot. Given the "vague links" between the tests themselves and what passed for general education, "how could you claim 'findings' or 'results' and link them to improvements?" asked Walker. "Why are we doing this testing?" Tenzer added. "Do we believe that resources will move around? Courses be ended?"

Assessment committee members not only heard questions and doubts from their faculty colleagues, they encoun-

*"... assessment is based on a fundamental misdiagnosis of the malaise of American higher education. Does anyone really believe that the failure of colleges and universities to produce adequately educated young people is the consequence of our failure to develop precise instruments to measure what we are doing?"*

—Jon Westling,  
Boston University, 1988.

*"Despite faculty apprehensions, there is no inherent need for assessment strategies to be crippling reductionist—excluding nuances and differences among students' patterns of development. Indeed... assessment need not be primarily quantifiable. It can and should bring into play the qualities of judgment and interpretation that faculty possess in abundance."*

—Carol Schneider,  
Association of American Colleges, 1988.

## The Mix of Methods

For many people "assessment" brings to mind "tests"; the two terms have been used synonymously at the K-12 level. Not surprisingly, then, many campuses that got into assessment three or four years ago (often with a tight, state-imposed timeline) turned first to the existing set of commercially available standardized tests.

Since then, a wave of test bashing has come and (pretty much) gone, with the result that standardized tests are still in use but typically as part of a larger package of "closer to the classroom" approaches: interviews, both individual and group; home-grown instruments tailored to a campus goals statement; depart-

mental capstone exercises or exams; portfolio analyses of student work over time; classroom research; faculty study of student transcripts, term papers, persistence rates, GRE and licensing-exam scores, and similar "indicators"; surveys of student satisfaction, motivation, and effort; alumni follow-up, done now with an eye to learning and its uses; case studies, ethnography, story collecting, and more.

Learning about learning isn't easy. What's needed, campus experience suggests, is a variety of methods, used over time, that give, together, the fullest, most accurate picture possible of student learning.

# The Perils of Gate Keeping

It's not outcomes assessment, but a cousin phenomenon: the external imposition of tests to regulate individual student progress—toward matriculant status, or into upper-division work, or for entrance into teacher-education programs. Texas, Georgia, Florida, and other states now deploy standardized tests of various types to enforce statewide standards for student progression. In Florida, all students at public institutions (or who receive state aid) must pass the College Level Academic Skills Test (CLAST) for entrance into upper-level coursework or for award of an AA degree.

Robert McCabe, president of Miami-Dade Community College, worries about such gate-keeping exams:

**"When standardized tests alone are used to make important decisions about individual students, and to control access, there is considerable prospect of damaging results. CLAST has done such damage in Florida.**

**"In my institution, half of our 50,000 credit students have a native language other than English. You**

tered frustrations of their own—and learned things. Velma Murry (family studies) reported her "distress and confusion" about attempts to design tests that fit the goals as written: "With the Culture and Modern Society exam, the committee wanted an exercise in abilities to reason and apply. What we got back from the work group seemed like a test of factual knowledge. Another test, in social science, came back all true and false." To the surprise and dismay of engineers on the committee, an exam on science and technology had to be "dumbed down" so that a broader spectrum of students could even understand it. "The whole process made us begin to think about how we're contributing to the general education of our students within specific disciplines," Murry reported. "It made us realize we've all been off doing our own thing."

A more troubling aspect to the committee's work arose with the prospect of students actually taking the new tests in summer and fall of '89. Murry put the matter succinctly: "How can

can imagine the task of getting these students through the composition section of CLAST—particularly since it is timed and may be on a topic with which they have no familiarity. With the holistic scoring method that is employed, there is virtually no prospect for a second-language student to meet the 1991 standard. The situation makes it very tempting (I heard this proposed on one of our campuses) to turn the English curriculum into a program to teach students how to produce a credible essay within extreme time constraints—simply test preparation. But this flies in the face of what most faculty tell me they *should* be teaching students: to organize ideas, to outline, to revise, and to use dictionaries. The need to help students pass CLAST drives us toward a curriculum that the faculty do not support, that is less rich academically, and that is certainly not in the best interest of students.

"CLAST is having a particularly devastating impact on minorities. For instance, the number of blacks proceeding to upper-division work has already been cut in half; a study has shown that a substantial per-

centage of those being denied upper-division access would be successful. One must ask, in a country where a severe problem is clearly indicated by the small number of minorities advancing through each level of education, why would a program be utilized that eliminates many of those who have demonstrated success in the lower division? It simply doesn't make any sense. It is a terrible waste of human resources.

centage of those being denied upper-division access would be successful. One must ask, in a country where a severe problem is clearly indicated by the small number of minorities advancing through each level of education, why would a program be utilized that eliminates many of those who have demonstrated success in the lower division? It simply doesn't make any sense. It is a terrible waste of human resources.

"As the growing damage to the people and economy of the state becomes more apparent, political forces should be more willing to modify this program. I hope we can make some changes in the Florida program—not do away with it but make it beneficial to students. We need to keep in view two things that distinguish American higher education from other systems. One is institutional and curricular diversity, which is threatened by the imposition of a standardized test; and the other is giving second and third chances to students, which is threatened by making judgments about students on the basis of a single criterion, a standardized test."

you ask a student questions about matters you've provided them no opportunity to learn?"

The goals statements themselves, six in all, typically begin with a general sentence followed by a dozen or so quite specific "the student should" statements, the latter a negotiated mix of the ideal, familiar, and hoped-for. Nowhere was it put down—this by design—that goals (and the instruments designed to assess them) had to reflect actual student courses of study; nor, given the looseness of the requirements, could they.

What the process of developing and administering assessment instruments uncovered—usefully, most would argue—were real disagreements about what should be taught. The thirteen "Student Goals for Foreign Language," reflecting the aims of faculty from other departments, emphasize speaking skills, cultural sensitivity, and a countering of ethnocentrism; many faculty within the language departments, on the other hand, consider it more important to teach grammar, vocabulary, and litera-

ture. "When new tests are rolled out reflecting wider viewpoints, chances are high they'll ask questions and turn up information some departments do not want to hear," Barbara Wright told us. "Assessment surfaces all these conflicts, like it or not. That's a not-so-secret reason for opposition to it."

Overcoming that opposition will take time, determination, and *allies*—one of the things we looked hard for. John Casteen III, president at the time of our visit, had declared his "pride" in general education at Connecticut but had not personally involved himself in the program or with assessment. Provost Thomas Tighe has been an enforcer of "publish or perish" in pursuit of status as "a major research university." The university senate and faculty union (AAUP) are deeply enmeshed in issues raised by funding shortfalls. Most faculty complaints about general education could no doubt be met through accommodations far short of those implied by assessment. Student voices are barely heard on the matter.

Meanwhile, though we heard little

about it from UConn faculty, the State of Connecticut has an assessment mandate in place, perceived as “permissive” but in fact stipulating a very full range of activity—in general education, in every major, even for graduate education—and it wants this all to happen soon. Arts and Sciences Dean Frank Vasington was visibly uncomfortable in describing this for us—he’s sat on the state committee for three years now—and wishes it would all slow down. Ruefully he acknowledged that most people on his campus have no idea yet about what the Department of Higher Education has in mind. “We have a big task ahead of us in educating faculty.”

Nevertheless, Vasington sees the committee’s work as “most valuable” in producing “interaction across departments and colleges, starting conversations about the links between teaching, curricula, and goals. We keep adding and adding courses without asking about coherence. . . . Assessment will turn up that issue. We impose requirements on students without knowing they do what we claim. Are we doing it [general education] right? We don’t have the evidence.”

Faculty involved in this venture seem determined to have that evidence, and to use it to raise larger issues about general education and the undergraduate experience. Their game plan, it seemed to us, is to begin with goals (in themselves hard to assail); test for their attainment; marshal evidence of success and shortfall; then, where necessary, push for corrective action. It’s a right plan, but not an easy one.

**O**ur visits to UVa and UConn came at what must be seen as early stages in their progress. The test of what they’ve done—and of their respective mandates—lies five or even ten years down the road. Even so, their early work is illustrative of assessment’s difficulties and its potential. We see in both cases stories about questions that are unfamiliar and, in both the asking and answering, *hard*; they’re stories, too, about important gains to be made by engaging those questions. Stepping back from the two stories, we note several themes they share.

## James Madison University

Assessment that aims for a picture of the whole student, in and out of class, requires the involvement of parties from across the campus; some have called such assessment “the ultimate interdisciplinary challenge.” It’s no easy task, but James Madison University has made considerable progress with it.

“We’ve tried,” says JMU assessment coordinator Dary Erwin, “to use assessment to build a more connected picture of what happens to students.” This entails collaboration between student-affairs and academic faculty, inquiring together into “factors in the campus culture that enhance or inhibit student learning and development.”

- The fact of assessment at UVa and UConn has prompted important *new conversations*. What one sees is not that the entire institution embraces assessment, or through it the cause of undergraduate reform; rather, on each campus we discovered a significant group of faculty (and a few administrators) who found assessment’s questions intriguing and who wanted to address them—together. Along the way, they have noted an absence of collective conversation about larger aspects of student learning, and have fashioned places and ways for such conversation to go forward.

- Those conversations have, on both campuses, already had an important effect: They’ve forced attention to *issues of undergraduate purposes* and made clearer the need for systematic information about their achievement. UConn seems a step ahead in addressing stage-two issues of this process, that of the match between university goals and those of specific courses as taught; UVa’s initial thrust has been to learn more about the student experience of coursework in a context of wider university life. (Each is now coming to the other’s issues, UConn via focus-group interviews of students now underway, UVa through its new proposals for revised general education requirements.)

Whatever the differences in direction and pace, both campuses have progressed toward more intentional, evi-

dent, for example, in JMU’s approach to assessing general education. Faculty have designed instruments to look at knowledge that students develop through general-education coursework, but they’ve also focused on cross-cutting abilities (competence in using the library, and computing skills, for instance) and on areas of personal, psychosocial development. In the latter, especially, assessment has involved joint work by student-affairs and academic faculty to trace the effects of students’ out-of-class activities on overall learning and development.

dence-based thinking about undergraduate education. Seen this way, assessment becomes a powerful way to set in motion *processes* of thought and conversation—of clarifying aims, then of holding aims against what is achieved—that will be at least as important in the long run as anything learned from data.

- Both stories illustrate that *answering assessment’s questions is far from easy* but that progress can be made. Early on, both campuses learned (as have a hundred others) that there are few off-the-shelf ways of addressing the particularities of one’s own situation. Our story of the two skips over the long hours and semesters of work that went into each’s search for right methods, then into the construction and debugging of instruments devised by faculty—research-oriented faculty with high expectations about what constitutes credible evidence. Are the resulting instruments perfect? No, indeed. Will they become better over time and with use? Surely. What is their present value? They’ve put on the table important evidence where none existed before, suggestive if not conclusive, occasion enough for discussion about possible improvement.

- On both campuses we became more aware of the *complicating effects of state mandates*. Doubtless it’s true that the effort at UVa would not have occurred but for pressure from SCHEV. Nevertheless SCHEV’s sponsorship

*"The support of [assessment for accountability] by academic administrators is one more step in their metamorphosis from stewards to managers. The consequent diminution of the faculty's opportunity to exercise professional judgment would be one more step in their metamorphosis from professors to teaching assistants."*

—Ernst Benjamin,  
general secretary of the AAUP, 1988.

*"... it is inappropriate and unrealistic to expect professors to subordinate everything to helping students achieve a set of shared objectives. But it is equally wrong for faculties to pay no attention to common goals and to ignore the question of how well these aims are being realized."*

Derek Bok,  
writing in *Change*, 1986.

cast a shadow over the assessment venture at UVa. Campus parties repeatedly asked us to interpret for them the state's "real intentions here, what would satisfy them. . . ." Faculty members especially had only the most rudimentary (and sometimes dead wrong) sense of what the mandate actually says ("value-added tests, right?"). That context made it hard for faculty to feel ownership of the effort and for administrators to see how the state's questions might connect with institutional ones. Ironically, what SCHEV may have intended as "openness" to a variety of campus approaches came to be seen locally as mixed and confusing signals. (Indeed, the difficulties of communication here—from legislature, to state board, to agency employee, to campus administrator, to faculty—are mind-boggling in the ways they can, and did, go wrong.) Meanwhile, in Connecticut, the state's mandate kicks in more fully only this year, its impacts on campus work as yet unknown. Its positive reception seemed less than assured to us last fall.

• Both stories bring into view the *culture necessary* for all of this to go forward, and the forces working against it. We see assessment asking questions about undergraduate education in places where that function may be only the third or fourth most important thing going on. Particularly in the case of general education, it's asking questions about a function that no one seems to own or get paid to worry about. Assessment wants conversation about collective responsibility for general, cross-cutting purposes in organizations that are highly decentralized, that prize individuality and autonomous work, and in which specialization reigns. It asks questions about teaching and its effects in environments in which many faculty don't think of themselves (first) as teachers, and where reward structures lead elsewhere.

Discouraging as these points may seem, they are not the whole story. Indeed, they make clearer the significance of accomplishments at UVa and UConn. A larger point, though, remains: the questions of assessment, and the responses they would evoke,

find on many campuses today little tradition or clear reward.

## V. Against the Grain

Assessment has had a difficult start on *many* of the campuses we've visited in the last four years. What we've come to understand is that the questions it poses run up against some of the deepest structures and habits of academic life. Four factors in particular make assessment "against the grain."

**First**, assessment runs against the grain in asserting faculty and institutional responsibility for student learning. On many campuses the tradition is otherwise: Teaching is understood to require mastery of subject matter, clear delivery, prompt and fair grading, the keeping of office hours; as for learning, "the good students will get it." Assessment, on the other hand, assumes that the point, indeed the test, of good teaching is student learning—and that there's a shared responsibility (with the student) to make that learning happen.

To say that faculty (and institutions) have a responsibility for student learning sounds right enough, and in many community colleges and liberal arts institutions you'd get ready agreement. Even on the most research-oriented campuses, one finds *many* faculty who care about students and work hard at their teaching. But there's a difference between that admirable disposition and believing, in the spirit of assessment, that "the test of teaching is learning."

Moreover, even where the rightness of that test is agreed upon, what does it really *mean*? It's one thing to say, yes, student learning is my job; quite another to act on that view on a daily basis. How, faculty members wonder, do I get clear about the outcomes of teaching? How do I know and ensure their achievement? What kind of evidence "counts"? How does this change what I do in my classes?

**Second**, for many faculty, the tasks and processes assessment would evoke threaten sacred territory. One of the reasons people choose faculty careers is because they prize autonomy: the freedom to pursue one's own research agenda, to teach what one wants (behind closed doors no less), to set one's

## Alverno College

Assessment at Alverno College dates back to 1973, when faculty and administrators reshaped the curriculum around eight cross-cutting abilities, a transformation that led in turn to a search for new ways to elicit demonstration of those abilities from the student. The faculty developed a set of principles and practices that would constitute a framework for performance assessment; those in turn led them to the assessment-center method of business and industry, which faculty adapted to educational purposes by focusing not on selection but on diagnosis and development.

Today Alverno's assessment program encompasses scores of different activities used to monitor and foster on-going, individual student learning. It's best understood not in terms of methods (there are many) but as a set of educational principles:

—assessment puts the focus on the student's progress toward publicly identified learning outcomes, with explicit criteria for success.

—assessment calls for the student to integrate what she knows with what she can do—the concept of *performance*.

—assessment is not an "add-on" but an on-going, integral part of the

own hours, and so on. Such autonomy occasionally becomes a source of complaint by external parties against "free-booting academics," but it's critical to scholarly inquiry and to the special relationship between teacher and student that is rightly guarded.

It is against this high regard for individual autonomy that assessment interposes questions about a *collective* faculty responsibility for student learning. It says to faculty, "Your job is not only to tend to the learning of your own students, but to worry about how that learning relates to other courses, and to ask what students' learning over many courses adds up to." Nowhere is this expectancy more against the grain than in respect to general education, where the absence of collective attention to coherence and outcome has so often resulted in what Ernest Boyer years ago labeled a curricular "disaster area."

Within departments and majors,

learning process.

—assessment entails feedback to the student: detailed, behavioral diagnoses of strengths and weaknesses through which she eventually develops the skills of self-assessment necessary for independent learning.

—assessment entails externality; student performances are evaluated not only by their teachers but by external assessors.

—assessment samples student performance in a variety of settings (the major, support area, off-campus experiential learning, etc.) and in multiple modes (writing, speaking, group interactions, etc.).

Assessment at Alverno focuses on the individual student. But to pursue larger questions about impact and effectiveness, the college also employs sophisticated program evaluation. In Alverno's Office of Research and Evaluation, funded out of the college budget and in place since 1976, researchers examine the impact, value, validity, and effectiveness of Alverno's educational assumptions and programs, and work with faculty to refine the links between teaching, assessment practice, and long-term learning outcomes.

too, assessment calls for faculty to get a lot clearer among themselves (and with students) about collective aims for cumulative learning. The language here—of "goals," "objectives," and "outcomes"—is enough to put off many faculty. Even more problematic for them is that assessment doesn't stop with goals; it wants to ask about the connectedness of goals to particular courses and instruction. Pretty soon, as one faculty member told us, "It looks a lot like assessment will be telling me how to teach my course. What ever happened to academic freedom?"

Marching everyone to the same music is *not* what assessment's advocates have in mind; we've not found anyone arguing in its name for the lockstep, teacher-proof curriculum of the high school. Nor is academic freedom, except in its mushiest and most sweeping sense (allowing anything to go forward in any classroom), at issue here. Assessment does, however, pose a coun-

*"If there is one generalization that can be made about the culture of the academic community, it is that unusual value is placed on acquiring and using information. To refuse, in principle, to investigate the results of our own actions because such an inquiry is deemed somehow inappropriate is to display inconsistency of the highest order—a fact not lost upon those outside higher education. Like any investigation, assessment involves vast technical difficulties and undreamed-of subtleties. However, difficulties and subtleties have hardly stopped us in other areas."*

Peter T. Ewell,  
writing in *Change*, 1985.

ter to faculty autonomy by calling for work together toward common goals for which *some* shared responsibility is assumed. It need hardly be said that on many campuses that's asking a lot.

Third, and a next "against the grain" circumstance, is that prevailing reward structures make it difficult to take assessment and its aims seriously.

For faculty on many campuses, assessment's questions arrive at a time of rising pressure for research productivity and publication. Insofar as assessment is about better teaching, the climate for it is distinctly uncongenial. Many assessment practitioners would agree (at least privately) with what we heard from Christine Young of the Academic Search Consultation Service: "Until evidence of teaching effectiveness is taken seriously as a criterion for hiring, promotion, tenure, and merit, those faculty who take teaching (and assessment) seriously may continue to function at the margins."

*"If there's one thing social science research has found consistently and unambiguously... it's that people will do more of whatever they are evaluated on doing. What is measured will increase, and what is not measured will decrease. That's why assessment is such a powerful activity. It cannot only measure, but change reality."*

—Linda Darling-Hammond,  
Rand Corporation, 1988.

Young's comment points particularly toward research universities (and aspirants to that status). But even on the larger number of campuses where teaching is the primary mission, it's often not clear that work on *assessment* will be rewarded. A relevant fact here is that many "teaching institutions" are characterized not by the presence of special seriousness about student learning but by the absence of research expectations; even on those campuses, then, rewards for assessment are in doubt. Faculty naturally ask themselves: If I do this difficult thing, will anybody care? Where's the payoff? Will it ever make a difference? What will I have to give up to do it? Are the costs in time and frustration worth it?

These questions become especially problematic in that assessment presumes collective activity. Where and how does *that* get rewarded? At Rhode Island College (with assistance from FIPSE), departments that demonstrate

gains in student learning receive modest but significant rewards such as increased travel funds or new equipment. But that effort is the proverbial exception proving the rule—that reward structures in higher education almost universally run to the individual, and therefore counter to the aims of assessment.

Rewards for assessment (and for student learning) are problematic not only for individual faculty but for institutions. Where assessment is mandated, it's often unclear exactly what the state wants, even what priority it attaches to assessment. As a Big Ten chancellor told us, "The list of things the state wants us to pay more attention to gets longer all the time—school improvement, research parks, medical services, waste disposal. . . . It's hard to know what kind of attention to give assessment, and how it will pay off."

Assessment in many states, in fact, comes as the latest item on an already crowded "quality control" docket—program review, special audits, accreditation, certifying exams—none of which will be eliminated to make room for assessment, and many of which include information that might be (but usually isn't) treated as relevant to assessment's questions. Indeed, much of the information generated by existing mechanisms is ignored at the state level, campuses believe, raising concerns that assessment will go down a similar road, lost in a continuing din of state "initiatives," with no tangible reward for the institution that resolves to do assessment well.

And speaking of rewards, the wary ask, how many states have come forward to fund the *improvements* that assessment identifies as possible or necessary?

Public institutions have come to know well the bases on which they will be rewarded—for sponsored research, victories on the football field, and sheer enrollment. But as Missouri's higher education commissioner Charles McClain observed last fall at a meeting in Santa Fe, "How often has an institution been rewarded because its students *learned more*?"

**Fourth**, the assessment movement itself seems to embody a growing tension between campus and statehouse, rais-

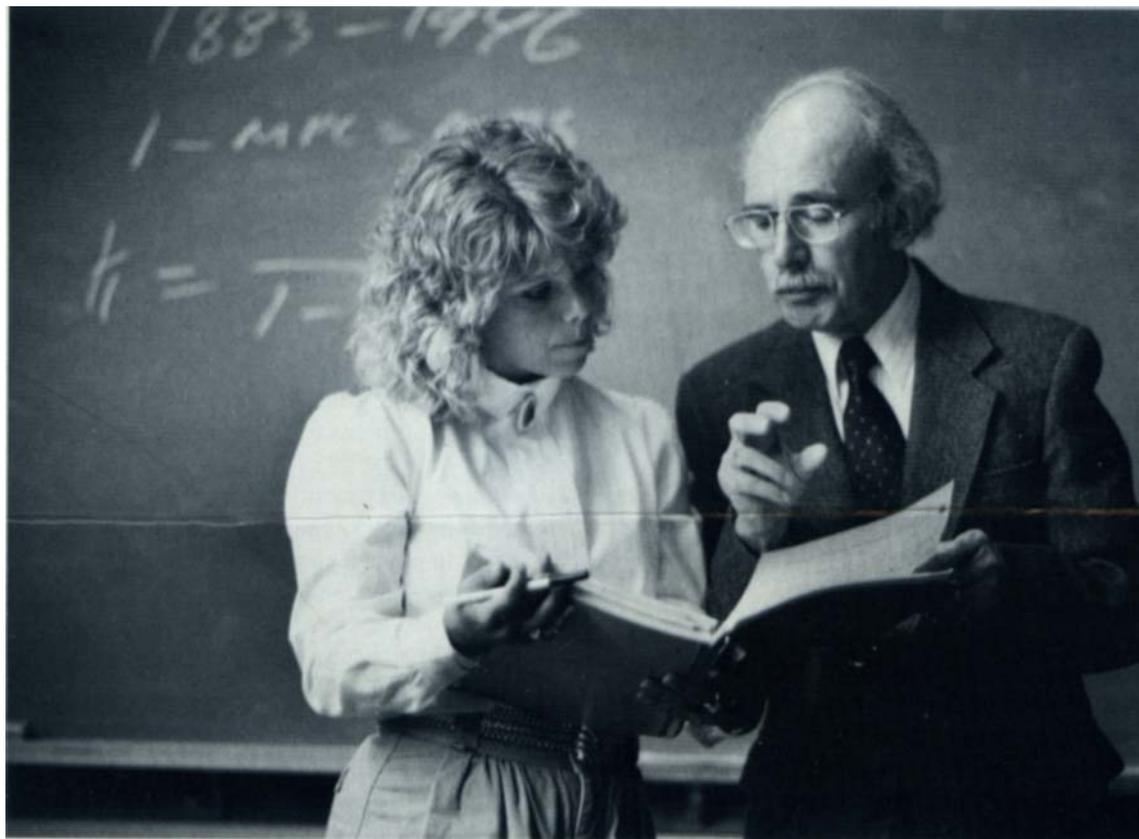
ing questions about "whose agenda" higher education will serve.

On the one hand the states complain about their colleges and universities: that they pay too little heed to public agendas (like school reform); that academics are off on their own agendas of institutional and personal aggrandizement; that tuitions keep going up but that student attainment and performance levels do not; that faculty are coddled, teach too few hours, are never in their offices, and leave the important work of educating undergraduates to foreign TAs; that higher education has been off the hook too long when it comes to providing evidence of effectiveness. One recognizes the voices here not only of legislators and governors but of William Bennett and Charles ("Profscam") Sykes.

On the other hand, institutions look around and see signs of threat and intrusion on all sides: states lurching from agenda to agenda; mandates rolled out one after another, hastily enacted and poorly understood; unreal timelines for "results"; expectations that colleges will overcome, with no increment of resources, years of public neglect of elementary and secondary education; and, more generally, that not much new money stands behind all the talk about the new importance of education.

The two "sides" here are obviously overdrawn; these arguments have a long history and take very different form in different states. But the heat behind them is a fact and at a high point right now; such tensions are an essential backdrop to much of what's going on today in the name of assessment. They explain in part why there are mandates, and why those mandates are not always eagerly embraced.

The good news is that work is proceeding on several fronts to build better state-campus understandings about assessment. With FIPSE support, the Education Commission of the States and the National Governors' Association are embarked on a two-year program of "seminars" that bring to the table campus and state officers for frank exchange about these matters, and to hash out the common ground behind assessment's questions through position papers and policy recommendations.



*UTK's assessment program emphasizes the use of assessment results in planning, budgeting, and program review.*

## VI. Two More Campuses

The ability of a state or institution to work through assessment's questions, to understand them and find useful ways to engage them, depends on several factors, among them time. This is especially clear when you look at the handful of now-mature assessment programs. It was to two such institutions that we turned in order to understand assessment's longer run: the University of Tennessee at Knoxville, a flagship university (like UVa and UConn), and King's College, a private liberal arts college in Pennsylvania. There are great differences between the two and in the ways they understand assessment, but from both come stories about the institutional rooting of assessment and the circumstances under which that rooting can occur—and make a difference.

### *The University of Tennessee at Knoxville*

To anyone strolling the 288-acre Knoxville campus of the University of Tennessee, the complexity of the place is palpable: 15 different colleges and

schools, over 70 departments, 19,500 undergraduates and 5,500 graduate students. A huge, brightly painted football stadium dominates the central part of the campus, whose latest jewel is a recently completed, state-of-the-art library. As a land-grant institution, UTK juggles diverse missions of research, teaching, and public service to the citizens of its state. There is, one quickly understands, no single university entity—no *one* UTK education—and asking, as assessment does, “what

it adds up to” is a tricky proposition. Given that context, UTK has fashioned an assessment program of many parts, with a locus of activity at the departmental level, and with strong emphasis on the use of assessment results in planning, budgeting, and program review.

The character of assessment at UTK is best understood in the context of state-level activity. In 1979 the Tennessee Higher Education Commission initiated a unique system of “perform-

## Administrators and Faculty—Communicating?

Several times each of us has had the experience of being invited to a campus by an administrator, warned in advance that “faculty here will never buy assessment.” But in meeting with faculty we find a great many taken by the questions behind assessment and by the intriguing ways it can inform their own concerns about student learning. At day's end, the skeptics remain the administrators . . . whose deeper worries often turn on past misuses of data by a board or the press.

“The beauty of assessment,” more than one dean has told us privately, “is that it's the best prompt in years for faculty development—a term I can't use out loud here.” “The best thing about assessment,” a faculty committee told us last spring, “is that finally we found something that got administrators to sit down and talk with us about students and the support we need to get a better handle on their learning.”

# International Developments

**The same economic and political pressures that have given rise to an assessment movement in the United States have produced parallel developments around the world—assessment as a governmental device to prompt higher institutional and student performance—with a host of familiar arguments thereby raised about accountability vs. improvement, qualitative vs. quantitative methods, externality vs. faculty ownership, and so on. Assessment developments are especially far along in the United Kingdom,**

**Australia, the Netherlands, and China.**

**In late July, sixty delegates from twenty countries convened at Scotland's St. Andrews University for the Second International Conference on Assessing Quality in Higher Education. In early August, over 300 Chinese educators (and eight Americans) gathered in Beijing, Wuhan, and Shanghai for a Sino-American Symposium on Assessment and Evaluation in Higher Education.**

ance funding," whereby a portion of state funds for higher education is based not on enrollment but on "performance criteria," defined today as: 1) the percentage of programs eligible for accreditation that are accredited; 2) student achievement in general education, including a "value-added" dimension; 3) student performance in the major field; 4) graduates' satisfaction with the educational experience; and 5) the existence of a plan for using the results of assessment to improve educational programs.

Institutional attention to these criteria was initially voluntary, but big bucks made them hard to ignore. Campuses undertaking assessment and compiling relevant results in an annual report became eligible to receive bonus funding—originally 2 percent of the instructional component of the education and general budget for each institution, and then, as of 1984, 5 percent, and currently 5.45 percent. For UTK this now means more than \$5 million a year.

After a decade of development, the scope of assessment at UTK is impressive. To measure student achievement in general education, the university adopted the ACT-COMP exam, given to a sample of freshmen and seniors in 1983 and used in 1988 to calculate gain scores. Since 1985-86 all graduating seniors have been required to take the exam. In 1989, for the seventh year in a row, students scored above the national norm for that exam.

To measure outcomes in major and professional fields, a broad array of methods is employed. Of UTK's more than 100 academic programs, approximately half have adopted an available standardized test—the GRE or a licensure exam, for instance. But some 47 programs have designed their own methods, ranging from multiple-choice exams to performance-based simulations and projects in senior capstone courses.

Finally, in order to learn more about student satisfaction, UTK employs a variety of annual surveys, designed by faculty specialists in survey research. Survey populations include current students, dropouts, and alumni; this year information about employer satisfaction is also being gathered.

As at UVa and UConn, assessment at UTK did not unfold without an occasional sticking point. Even today, faculty are not all enthusiastic about the university's assessment emphasis. As coordinator of the assessment program, Trudy Banta has heard her share of complaints and concerns—many predictable and appropriate. The ACT-COMP and subsequently the ETS Academic Profile (which the institution pilot tested in 1987) have turned out not to be a good match for the general education program at UTK (faculty estimate the match at 29 and 30 percent respectively). It's proved difficult as well to connect test results to anything specific that might be fixed. More recently developed tests (CAAP and Col-

lege-BASE) have been tried, to similar effect. Faculty-designed exams, while they promote good conversation and faculty ownership, have not been totally successful at measuring higher levels of cognitive functioning.

There have, in short, been plenty of problems, but over time assessment has won friends as well. As Banta says, "Many faculty still question the purposes of assessment, but in each department you find people interested in teaching and learning who've used it as an impetus to do things they wanted to do."

Assessment has survived its pitfalls in part because of Banta herself. An important part of the UTK story is her tact, her ability to help faculty see the good in assessment, her knack for translating results in ways that make sense to various audiences.

But assessment's progress at UTK has been a consequence of more than personality. From the start, the venture has enjoyed top-level *institutional* commitment. As a result, assessment has been built into on-going institutional processes and into visible mechanisms of decision-making that faculty understand and value.

One such mechanism is program review. Since 1983 departments have been required to add to traditional measures of effectiveness (resources, faculty credentials, ability of incoming students, and the like) "information about instructional objectives" and "evidence of effectiveness" in helping students attain them. Today, program reviews at UTK routinely include data from student surveys and from relevant exams in general education and the major. This link "was one of the most productive we've made," says Banta. "When faculty must do program review, and the spotlight is on the department's goals and effectiveness, they want to make the best case possible. . . . They *want* that information on student learning."

Assessment has also become a routine part of UTK's annual planning process, providing data relevant to about 40 percent of the institution's goals. One such goal entails a training program for TAs, the design and implementation of which depended on information from student surveys; the

value of that venture was confirmed by subsequent surveys, which showed student satisfaction with TAs markedly increased.

The ultimate rooting of assessment took place at the system level. In 1988, Lamar Alexander, the governor when performance funding was initiated, became chancellor of the UT system. Five months into his tenure he scrapped an existing administrative structure and replaced it with two new positions: a senior vice president in charge of planning, Homer Fisher, and a vice president for assessment, Michael Nettles. "Homer's job is to ask where we're going," Nettles says. "Mine is to find out if we're getting there."

Is the university getting where it wants to go? There are many reasons to answer yes; progress has been made at many levels at UTK, with assessment a part of the story.

Some gains came early and with relative ease. Departments in the college of agriculture, for example, were among those that designed their own exams. In the process of doing so—long before they had any results—they saw the value of being clearer about outcomes and set about devising student learning goals at the departmental level, then for individual courses.

Not surprisingly, many other changes have taken time. The geography department began using its own comprehensive exam for seniors in 1983; six years later, faculty came forward with a new curriculum that reflected what they had discovered about their students' learning from that exam.

Some effects of assessment are harder to point to but, as we heard, equally important. Dorothy Habel, an associate professor of art history, tells the story of stormy discussions with her colleagues, who were predictably disturbed by the notion of a state-mandated test. Nevertheless, it was in discussing what such a test might look like that the department "first grappled with what we expected of students." "It was the first time we sat down as a department and talked about what we do in class." Similarly, in the department of food technology, the entire faculty took its own departmentally designed exam, with the consequence, says chair Hugh Jaynes, that everyone

## Kean College of New Jersey

Several years ago, when Kean accepted a \$3.8-million challenge grant from the state—a big part it earmarked for assessment—an eight-month debate ensued about "how to do this thing right." The product, in 1986, was an eight-point statement of principles, fully subscribed to by the college's president, senate, and powerful union. Thus endorsed, assessment became the spark for four years of creativity and change at this 13,000-student state college.

The Kean model eschews external reporting and the assessment of individual students; it minimizes institution-wide measures and focuses on self-assessment within majors and programs, some forty in all, with each faculty unit at liberty to devise its own approach—no two of which, we found on a visit last January, are the same. In many departments we heard tales of early foot dragging ("We thought we'd just give a test"), of challenges from colleagues ("Don't be superficial! Where's your integrity?"), then of goals getting clarified, courses realigned, teaching altered ("It took a



mountain of conversation")—all this before any data entered the picture.

The results? By faculty report, assessment has ushered in a time of greater communication, trust, introspection, and fact-based decision making. Student retention is on the rise. The next step at Kean, it seems, will be clearer, higher, more consistent standards for student performance. "Assessment changed the way we think about students," one chair told us; "They may be underprepared, but now we know—we believe—they can learn."

understands the total curriculum more clearly.

"Bad tidings" have sometimes turned out to be most useful. A first round of COMP testing showed low scores on "problem solving" in certain colleges; faculty were predictably upset. But an analysis of their own classroom exams showed that only 15 percent required students to solve problems. "Then they wanted to know more about how to teach and examine for that ability," says Banta. A consequence was faculty proposals for experiments in four colleges to increase problem-solving and critical-thinking skills.

Results on student satisfaction surveys have been particularly effective at prompting change at UTK. Looking closely at the relationship between student satisfaction and the amount of faculty-student interaction, many academic units have taken steps to increase opportunities for such contact. The percentage of students who report they "know no faculty member well

enough to ask for a recommendation" has dropped from 47 percent in 1983 to 27 percent in 1989. Determined efforts by faculty and support staff to take seriously student reports about their freshman experience at UTK prompted a whole series of small changes, which raised over time the freshman-to-sophomore retention rate from 62 to 71 percent.

Assessment is never, on any campus, going to be an unmitigated success. Prowl around UKT and you'll find faculty who don't like it, even more whose daily work is barely touched by it. Nevertheless, assessment's questions have now worked their way into the mindset of the institution. Its processes have put new issues on the table and new kinds of information—about student learning—into budget and program review. For ten years now assessment has significantly supported the work of the scores of people at UTK who care about improving the undergraduate experience.

*King's sharp vision of undergraduate education, not a mandate, led to its "course-embedded" assessment program.*



### **King's College**

From some points of view, nothing could look more different from UTK than King's College. A private liberal arts college in Wilkes-Barre, Pennsylvania—the heart of coal country—King's was originally an all-male institution, enrolled its first fully coeducational class in 1970, and now has a full-time student body of 1,750, plus 600 part-time students, recruited mainly from Pennsylvania, New York, and New Jersey. The view book touts a "spirit of family," 34 degree programs, and a core curriculum "cited by several national academic associations as a model."

Integral to King's outcomes-oriented core is a comprehensive program of student assessment. Indeed, it's the college's sharp vision of undergraduate education—not a mandate—that led to its "course-embedded" assessment program: assessment that's located within courses, run by faculty, and focused on individual student learning. At King's, assessment has little to do with the kinds of program-effectiveness ques-

tions that drive work at UTK; it is first and foremost a *pedagogical* strategy.

The King's story begins in the late '70s—before the rash of reform reports and considerably before any "assessment movement"—when the college undertook a strategic planning process in which faculty asked themselves about "the proper definition of excellence." Their answer, writes Peter Ewell, was that "excellence necessarily rests not upon their selectivity but upon what they can demonstrably *do* with their students in the four or more years that they work with them."

For academic dean Donald Farmer, this new conception of quality took human form in a parting chat with a senior. Conversation took the usual turns. What was the student going to do after graduation? What classes had he particularly enjoyed? On the way out the door, the student paused and turned with "just one last thought. I think you should know," he confided, "there's a lot more teaching going on here than learning." Farmer heard him loud and clear.

It was learning that Farmer and the

King's faculty were after when they decided to take a closer look at student abilities to write. In 1982 the faculty began intensive discussions of writing across the curriculum. "It was through writing across the curriculum," English department chair George Hammerbacher recalls, "that we came to the outcomes question." As they worked their way into cross-cutting abilities (first writing, then critical thinking and problem solving), faculty confronted precisely the questions that lie behind assessment: about aims and purposes; larger, cumulative learning; shared responsibility for making that learning happen; and methods for knowing that it has occurred.

Answers to these questions can be found in the new core curriculum that faculty forged over several years and that went into effect in 1985. Required of all students and constituting half of the work needed for graduation, the King's core is a sequence of entirely new courses, each designed by a team of faculty who began not with questions of turf or coverage but by asking how each course would contribute to

student learning of eight “transferable liberal learning skills”: critical thinking, creative thinking and problem solving strategies, effective writing, effective oral communication, quantitative analysis, computer literacy, library and information technologies, and values awareness.

These eight outcomes were not only the framework for a new curriculum; they were the building blocks of King’s evolving assessment program. Rather than establishing a self-standing office responsible for assessment across the entire institution, King’s opted for what it calls “course-embedded assessment,” that is, assessment done by faculty with their own students in regular courses. In contrast to the practice elsewhere of episodic, large-scale testing, assessment at King’s is an ongoing process of monitoring individual student progress toward identified outcomes; its aim is to provide each student with feedback that promotes learning of larger, cross-cutting outcomes. Assessment exercises and exams are part of regular course activities, and, yes, they *count* toward the final grade. Beyond its “course-embedded” as-

pects, assessment at King’s is notable for its impact on faculty thinking. Associate professor of psychology Jean O’Brien remembers how she and her colleagues struggled to be clearer about purposes: “It really takes time to stand back and say what outcomes I expect and what criteria I’ll use to judge them,” she told us. Assessment made her realize a “need to be more explicit about what I’m looking for; the more explicit I can be, the more students learn.” She talks, too, about changing her teaching style—“I used to say I covered the material. Now I wonder if that wasn’t true”—and finds herself lecturing less and turning to more involving activities. She talks with colleagues in other departments and devises assignments that build on what students are doing in other classes.

What was striking in our conversation with O’Brien was that although “assessment” was what we were asking her about, most of her comments were about students and learning. And, indeed, assessment at King’s is hard to separate from the on-going processes of teaching and learning. Visitors to the campus often tell Farmer, “What

*“The biggest and most long-lasting reforms of undergraduate education will come when individual faculty or small groups of instructors adopt the view of themselves as reformers within their immediate sphere of influence, the classes they teach every day.”*

—K. Patricia Cross,  
University of California, Berkeley, 1990.

*“If we want to improve the quality of college education we should focus not on testing, but on better teaching.”*

—Ernest L. Boyer,  
interviewed by Ted Marchese,  
in *Change*, 1986.

## Seattle Central Community College

**Community colleges lead the way when it comes to entry-level assessment and the use of results for placement and advising of students. What has seemed harder for many of them is assessment of outcomes. With part-time, in-and-out enrollment patterns, asking assessment’s “what-it-adds-up-to” question is tricky. But at Seattle Central faculty are looking for answers.**

Work on outcomes assessment began several years ago, when a group of faculty read its way through the literature, then spent a year devising a set of outcomes for SCCC, including information-seeking abilities, multi-cultural literacy, critical thinking, responsible action, and a capacity for life-long learning.

“This is an evolving list,” says director of planning and research Jack Bautsch. “Our intent is not to come to final consensus but to keep the conversation going: to think harder about what we expect of students, how we help students meet those expectations, and whether we’re successful.”



How will the college answer these questions? In as many ways as possible, says Bautsch; and in ways that involve people across the campus. A faculty member in English is now pilot-testing a course in which students learn to assess their own learning. Many faculty have been trained to use K. Patricia Cross’s techniques of classroom research. Teams of faculty and representatives from industry are rethinking vocational curricula—their content,

pedagogy, and assessment.

In the next year, with a small grant from the Washington State Board for Community College Education, Seattle Central will work with two sister campuses to take yet another assessment step: a study of knowledge, abilities, and attitudes of students who transfer and complete the bachelor’s degree—information anticipating a district-wide consensus on student outcomes for the transfer program.

you're doing is just common sense." As one hears faculty talk about being more explicit with students, about clarifying what is expected and holding students to it, one is tempted to say, "That's what *all* good teachers do. It's nice, but it's not news."

There *is* a certain modesty to it; Farmer once described assessment at King's as "100 small experiments." But what looks like common sense was, says Jean O'Brien, nevertheless "a huge change"—a change best understood, perhaps, in terms of commitment and values: a commitment by faculty to teach for and gather evidence about a set of agreed-upon purposes, enacting "a belief that something better can happen for students."

"And does it?," we asked. What evidence is there that assessment at King's brings about that "something better"? Because the college's emphasis is on "what is adds up to" for individual students, little attention has gone to generating evidence of aggregate gains. But, if pressed, Farmer can indeed cite rising LSAT scores, increased library use, and attendance at on-campus cultural events. He notes, too, a rise in King's graduation rates that many campuses would envy: up from 58 percent six years ago to 71 percent today. Hardly modest.

What else? Asked whether assessment is "worth it," Farmer points to lunchroom conversations among faculty: "They don't talk about parking lots, they talk about student learning."

**K**ing's and UTK differ greatly in their approaches to assessment: UTK, a large public research university, operating in the context of a state mandate, uses large-scale, institution-wide methods, with a centralized Center for Assessment Research and Development; King's, a small, private college with no central assessment office but "100 small experiments," uses classroom-based approaches aimed at individual students. At the same time, both are stories about constructive ways of engaging assessment's questions; about the development over time of habits of inquiry about student learning; and about the use of information to improve teaching, curricula,

and services to students. More particularly, we think their stories illustrate four larger themes.

- We said it before and note it here again: assessment takes time. At both King's and UTK, developments took place over a period of 10 years. A visit to either campus six or eight years ago would have turned up much of the same faltering and uncertainty that characterizes the newer work at UConn and UVa. Good things come slowly: UTK's new geography curriculum took six years to emerge; the King's faculty devoted five years to curricular work before assessment entered the picture. These changes may not reflect the timetables that legislators and governors have in mind, but they're ones that matter.

- To endure, assessment needs to find a home within ongoing institutional mechanisms. At King's this meant building assessment into that most fundamental of institutional functions, the teaching and taking of courses. At UTK, assessment took legitimacy as it was built into program review and budgeting decisions. In both cases assessment developed not as "a train on its own track"—as Peter Ewell has found it to be in many settings—but in ways that become integral, as a means toward valued ends.

- On both campuses—as at UVa and UConn—the questioning and conversations prompted by assessment proved valuable in and of themselves. At UTK and King's, the value of these processes was recognized from the start. Faculty were given the time, space, and support to understand assessment and address its questions in their own terms. Faculty knew, too, that these questions were important to the institution and that time spent engaging them would be rewarded.

- On both campuses assessment found the leadership it needed. Top administrators sent consistent signals that assessment mattered; they provided resources and support—and kept pushing. Equally important was leadership from within the faculty, the presence of "idea champions" who took the time to learn, try new things, and speak out. *Both* kinds of leadership were necessary to what happened.

## VII. So What?

Few, *very* few, campuses have been at assessment as long and as successfully as UTK and King's. What, then, about the many, more recent actors in the assessment story? For how many of the eight out of ten institutions now reporting "assessment activities underway" can assessment be said to be making a difference? Five years into this movement, what do we know about the acid question: Does assessment improve student learning?

A partial answer comes from Peter Ewell, who has almost certainly made more assessment-related visits to campuses and capitols than any person in the country. Ewell's current estimate: in mandate states about 15 percent of campuses are moving ahead in ways that are likely to make a real difference; a second 15 percent are totally at sea—or stonewalling; the big group in the middle is complying as best it can.

The two of us have also been trying to answer the "so what" question—mostly by asking it of the dozens of campuses that are now three-to-five years down the road. What we've heard by way of response are a hundred stories—stories about the redesign of courses in a nursing sequence, the repair of an advising system, increased retention of minority students, a new goals statement for an engineering program, the more systematic use of student study groups in large science courses. . . . We've listened hard to these accounts of improvements and been excited by many of them. Whether they *prove* that assessment causes more student learning is another matter.

One circumstance is that questions about assessment's impact—its power to improve things—are still premature on lots of campuses. Programs (as we saw at UTK and King's) take time to develop and show results; those one-to-three-years old just won't have much to show yet. We asked participants at AAHE's 1989 assessment conference to tell us, as part of a conference evaluation, "What has happened on your campus as a *result* of assessment?" More than three-quarters of respondents passed over the question or reported that they "hadn't gotten to that point."

Even if they had, however, the possibility of proving a cause-and-effect relationship between assessment and improved learning is likely to remain elusive. In the messy, real world of campus assessment endeavors, the experimental conditions needed to attribute causation will never exist.

Leaving aside questions of strict cause and effect, one can nevertheless point to institutions that are active in assessment *and* have improvements to report. One such is Alverno College, where a highly sophisticated longitudinal study follows students through college and beyond. Nowhere, perhaps, are there better, more complete data about gains in student learning over time—data that would seem to document the impact of assessment. But behind those gains lies not only assessment (as most people use the word) but Alverno's powerful learning culture: consistency and clarity of purpose, teaching aimed at those purposes, a sense of responsibility to students, a sophisticated, institution-wide conversation about learning, and a view of teaching as a valued professional activity. Assessment at Alverno is essential to all of the above, but it's also part of a bigger picture, and separating out its effects is not possible—or particularly useful.

What one sees at Alverno and on other assessment-minded campuses—

at King's, UTK, Miami-Dade Community College, Northeast Missouri State—is that assessment is but one of three or five or eight major things going on that add up to high-impact undergraduate education. It's not only that assessment's effects can't be separated out, then; multiple, linked lines of work have to be in place for significant gains in learning to occur.

One sees a parallel phenomenon (or misconception) at the state level. Something you'll never see is a case where 1) a state enacts a mandate, 2) institutions "do assessment," and then 3) improvements ensue—though that's not far from the hope that has driven some state initiatives. This isn't to say that states shouldn't be patrons of assessment or that a mandate can't prompt good things. What it does say, and what ECS/AAHE case studies of eight assessment-active states have shown, is that a mandate alone—like assessment alone—won't get you far.

New Jersey's much-touted basic skills program is a case in point. Assessment—in this instance an entry-level, diagnostic exam sponsored by the state—played a key role in tackling the problem of student underpreparation. Assessment's power was in supplying numbers to describe the problem and to make it public in a way that became hard to ignore. But assessment was

*"Educators have to help design assessment systems that meet their own highest standards. Assessment has to ask and answer real questions. It has to rest on values that presidents, faculty, and trustees believe in and will live by in their work. It has to have integrity."*

—Thomas H. Kean,  
writing as governor of New Jersey,  
in *Change*, 1987.

*"The outcomes an institution looks for—and the way to go about assessing them—reflect in a way few other actions do just what the institution believes its role is."*

—Patricia H. Murrell,  
Memphis State University, 1987.

## Northeast Missouri State University

When Charles McClain, president of Northeast Missouri State University from 1970–89, talks about assessment there, he talks mostly about vision, mission, and the institution's sense of itself. He quotes from Peter Ewell's *The Self-Regarding Institution* about "the importance of an institution's ability to state clearly its particular mission and the particular types of students it intends to produce."

It is in fact a vision of NEMO and what it could be that has driven its evolving assessment effort. Three parts of that effort have been in place since the mid-'70s: 1) a *value-added* component, which seeks to measure student growth in general knowledge from the freshman through the sophomore year via tests from ACT; 2) a *comparative* component, which seeks to demon-

strate student achievement in the major field by means of nationally standardized tests, or, where no such test exists, a locally developed exam; and 3) an *attitudinal* component, which seeks to determine student perceptions of their own growth and satisfaction with the university and its services. In more recent years, NEMO has turned from heavy dependence on testing to a panoply of methods, including more qualitative approaches such as portfolios.

What hasn't changed over the years is an underlying commitment to questions about quality and to communicating that commitment (and assessment results) to the public. That strategy has paid off in big ways—in an expanding pool of applicants, increased state support, and high faculty morale.

*“Assessment—even now in its infancy—is already doing more for education, for institutions, and for faculty members than any other development in recent history.”*

—James H. Daughdrill,  
president of Rhodes College, 1988.

*“The democratic and humane values of our society must drive the purposes of assessment. [That means assessment that will] enable and empower more students, particularly minority and poor students, to get more and better education.”*

—Donald M. Stewart,  
president, The College Board, 1987.

only part of a larger program of improvement launched by people committed to making a change—faculty who rallied around, began to figure out why students weren't doing better, then worked backwards from test results to curricular and instructional matters, first with cooperating schools, then at the college level. Over the '80s, as follow-up measures have shown, major gains in student attainment have been realized.

The New Jersey example reminds us that assessment doesn't solve problems—people do, in this case (necessarily) people with resources. Assessment then comes back into the picture to tell those people whether their efforts are making a difference.

And so if the question is whether assessment fixes things, the answer is no. But what if you ask, Are things more likely to improve with assessment than without it? Does assessment contribute to a set of conditions where improvement is more likely? Our answer is yes. The case studies in this report and the dozens of other campus stories we've studied show that assessment can a) raise questions and prompt processes that help clarify collective goals and expectations; b) identify problems and put them “on the table” in ways that force attention; c) build habits of inquiry and a culture of evidence about student learning.

With the right leadership and sense of purpose, assessment can do these things even where it is “against the grain”—indeed, over time, it can help change the deeper institutional values and shape cultures more conducive to teaching and learning.

A final take on the “so what” question is to ask: Would we lose anything if assessment went away? A look at individual campus stories and at the larger movement suggests we would indeed. Assessment as a national phenomenon has captured the attention of an incredible number of people across a broad range of constituencies, from junior professors to governors; it has those people asking at least some of the right questions about teaching and learning. It has become, we think, the arena for the day's best discussions about undergraduate education and the most promising focus we have for

improvements.

In sum, where assessment is working, it's almost impossible to pull out the causes and effects. Where it works, it works because it's integral, not as a separate function off by itself but as a process woven into daily activity. Seen this way, assessment becomes a powerful, if insufficient, condition for change.

## VIII. Assessment and Change: Prospects for the Long Term

Barbara Wright recalls the moment in her work as director of UConn's assessment project when her FIPSE program officer wondered whether assessment in Storrs would be a mere “blip on the screen” or lead to “real change in campus culture.” “My God,” Wright panicked, “is *that* what we're supposed to do? Change the campus culture?”

Our answer—the answer implied by the story of assessment as it's told here—is yes; assessment *is* a story about institutional change. As such, it invites a final question: What, over the long term, are the prospects here for deeper, lasting change?

The answer to *that* question, we've come to believe, may have only a little to do with assessment itself. Institutions, given the day's pressures and mandates, will *do* assessment; the issue is, to what effect? And effect, as we think the stories in this piece suggest, has less to do with a technically correct doing of assessment than with a larger mindset that undergraduate improvement is possible, necessary, and a priority. It is a mindset about *quality*, one that assessment both presumes and can help to prompt, where quality is seen as a function of student learning and the improvement of such learning is constantly pursued.

Looking at institutions where assessment has had a hard time finding a home—where it's been particularly against the grain—one often sees good intentions and “right moves” but uncertain prospects for deeper effect. Why? Part of the reason would seem to lie in the absence of any widely felt need to improve. The institutional perception may be that “things are pretty good as they are,” or “if it ain't broke,

## Behind Outcomes

A major contribution of assessment is that it's made a much larger number of people aware of the importance of outcomes and of looking not just at resources and processes but at their effects (the "results" criterion). Interestingly, though, among the movement's more experienced practitioners, there's been a de-emphasis of late on outcomes, for a number of reasons.

An initial reason was that satisfactory measures of aggregate outcomes are few and far between; the summary measures that exist are often so general in nature that their "findings" are next to impossible to connect with specific, fixable features of teaching or curriculum.

A related insight from practice is that knowledge of outcomes may tell you how, at a given moment in time, you're doing at end-point, but that information alone seldom turns out to be helpful for improvement. What you need with it is context, the story over time, final effects laid against information about where students started and what happened to them along the way. That's why smart colleges are asking questions about student-faculty contact, time spent on studies, frequency and purpose of library use, and so on.

don't fix it," or, even if it is "broke," improvements are out of the question because resources are already so stretched. The problem here is not, of course, assessment. It is an attitude toward improvement—the lack of what we heard one campus administrator describe as "ya gotta wanna."

That's a good phrase, in fact, for the mindset we found at institutions where assessment has been embraced and made a difference. One thinks of the comment by Jean O'Brien at King's that assessment is "a belief that something better can happen for students." Or one thinks of Alverno, where assessment is best understood not in

## Where to Learn More

A few years ago there was virtually no literature on assessment, no place to turn for resources, no way to find colleagues. Now there are numerous routes to assistance.

If you have questions about choosing and using assessment instruments and methods, a good place to begin is The Center for Assessment Research and Development at the University of Tennessee, Knoxville. Director Trudy Banta and her colleague Gary Pike provide materials and training; contact them at 1819 Andy Holt Avenue, The University of Tennessee, Knoxville, TN 37996-4350.

The AAHE Assessment Forum was established in 1987 to advance thoughtful campus practice. Services include: assistance for campus practitioners, workshops and presentations, commissioned papers and conference presentations, national directories, an in-house assessment resource library, access to a national network of assessment practitioners, and an annual conference. For further information, write: AAHE Assessment Forum, One Dupont Circle, Suite 600, Washington, DC 20036. Forum director is Barbara Wright.

Materials collected by the AAHE Assessment Forum—many of them fugitive (state documents, campus

terms of any particular method or approach but as the necessary adjunct to a larger goal, the life-long learning of every student.

This drive "to make something better happen" also struck us at Miami-Dade Community College, where there's been a constant, 15-year press to do better on behalf of all students—many of whom face long odds against success. Every semester, it seems, some new, campuswide innovation is put in place; the longer-term results (assessed through attainment and transfer rates and, yes, test scores) are striking. Not incidentally, this pioneering college maintains a 13-person research office, headed by a

reports, conference proceedings, and so forth)—are available through the Educational Resources Information Center (ERIC) system. Use the identifier "AAHE ASSESSMENT FORUM" to retrieve them.

The U.S. Department of Education's Office of Educational Research and Improvement has sponsored and published a most helpful collection of peer-reviewed essays, *Performance and Judgment: Essays on Principles and Practices in Higher Education Assessment* (edited by Clifford Adelman, 1988). Its contributors delve into assessment's deeper, more technical concerns (e.g., construct validity in questionnaires, reliability in performance assessment) and provide extensive bibliographies. Order (stock #065-000-00342-2) from the U.S. Government Printing Office, Washington, DC 20402; the price is \$15.

Jossey-Bass has come out with numerous quarterlies and books on assessment in recent years. Write for a listing (address below). Also from Jossey-Bass comes *Assessment Update*, a quarterly newsletter and the best thing around for keeping up with recent developments, opinion, and new publications. A year's subscription is \$60, from Jossey-Bass, 350 Sansome Street, San Francisco, CA 94104-1310. *Update* editor Trudy Banta can be reached at the address at the top of this box; she is "always on the lookout for new authors and contributions."

dean, constantly to tell itself, its faculty, and its students how they're doing.

As we thought about this mindset about improvement, we noted that many institutions in the '80s have indeed sought to raise their "quality"—by hiring "stars" and pressing faculty for research output, or by imposing new admissions and curricular requirements on students. Assessment institutions, too, have sought to raise their quality, but in a very different way and place: within the classroom. They've done that by imposing heightened demands on *themselves*—to help every student who comes to them for an education to succeed.

"Some good things may happen as a consequence of mandates, but ultimately quality in undergraduate education cannot be legislated, regulated, or inspected into existence. . . . Context and commitment describe what's needed for assessment to succeed on the campus level. These things have to come from within, not from a statehouse."

—Daniel Seymour,  
UCLA, 1989.

"SUNY needs to move forward on academic assessment not because we are forced to by state mandate but because we want to as responsible academics."

—Joseph C. Burke,  
Provost, State University of New York, 1988.

## The FIPSE Connection

With the exception of two addresses on the topic by William Bennett (in 1986), the sponsorship of several studies by its Office of Educational Research and Improvement, and its new rules for the recognition of accrediting agencies (1988), the U.S. Department of Education has tended to play (at best) a supporting role in the emergence of assessment. As a phenomenon, assessment is campus-level and state-specific; most of the Washington-based associations and press (including the *Chronicle of Higher Education*) have paid scant attention to it.

A notable exception to this is the

Fund for the Improvement of Post-secondary Education ("FIPSE"), an innovation-minded agency within the Department of Education. From its program budget of just over \$10 million a year, FIPSE has funded several dozen assessment-related projects since the mid-'80s, plus their evaluation, with the result that a large part of what the field knows about practice is the result of this one foundation's endeavor. Practitioners credit FIPSE's director, Charles Karelis, and dedicated program officers for staying with this issue and for shrewd grant making over the years.

A characteristic of this approach to quality relates to mission and institutional identity. We've named 10 or 15 "assessment institutions" in the course of this article; there are vast differences of approach among them—between King's and UTK, for example, or Alverno and Northeast Missouri State. But what each of them shares with the others is a sharp sense of who it is and what it wants for its students. Everything these places do is to get closer to that end. Each manifests a larger sense of *educational* purposefulness that connects everything from the individual course and admissions policy and trustee oversight to institutional mission. Assessment is powerful in such places because there's a larger principle of *integrity* at work.

. . . Integrity and *accountability*. Early on in the assessment movement accountability became a dirty word; it was what "they" wanted, where "they" were bureaucrats and politicians with little sense of the subtleties of higher education. This is a perception that's starting to change, but a legacy remains. Even today, most of the talk about accountability focuses on that which runs from the institution to the state—expressed in the form of new reporting requirements.

Institutions that are taking assessment seriously have a different slant on accountability. While recognizing obligations to external publics, they want to look inside to deeper-running responsibilities.

They think, for example, of the institution's obligation to its students. Accountability here means delivering an education equal to that promised in recruitment, and commensurate with the student's investment in it (not only of money but time and effort). Some campuses are beginning to think, as well, of students' accountability for their own learning, and are teaching students to "self-assess," to take responsibility, to ask the "what-it-adds-up-to" question of *themselves* as learners.

Most important to the quality mindset behind assessment is the accountability that educators have as professionals to each other . . . to deliver, in their teaching and related work, on mutually agreed upon purposes (this, the professional obligation that goes with the autonomy that faculty have traditionally enjoyed). In asking questions about educational purposes, assessment leaves room for any number of answers, but it does assume that these are important questions, to which there are *some* answers, and that those answers will be broadly shared and worked toward by every member of the campus community: faculty, staff, administrators, and students.

With these deeper forms of accountability understood and in place, institutional aims find expression in the work of all members of the campus community. Where that's the case, you want questions about those aims—and evidence of their accomplishment. You want assessment. □