DOCUMENT RESUME

ED 449 770 HE 033 782

AUTHOR	Cross, K. Patricia; Wiggins, Grant; Hutchings, Pat
TITLE	Assessment 1990: Understanding the Implications. The AAHE Assessment Forum. Conference Proceedings (5th, Washington, DC, June 27-30, 1990).
INSTITUTION	American Association for Higher Education, Washington, DC.
PUB DATE	1990-06-00
NOTE	49p.
AVAILABLE FROM	American Association for the Study of Higher Education, One Dupont Circle, Suite 600, Washington, DC 20036-1110 ($12). Tel: 202-293-6440; Fax: 202-293-0073.
PUB TYPE	Collected Works - Proceedings (021)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Accountability; *Educational Assessment; Educational Improvement; *Higher Education; *Learning; Liberal Arts; *Student Evaluation; Teaching Methods

ABSTRACT

Three papers from the American Association for Higher Education conference in 1990 are presented. The first, "Streams of Thought about Assessment" by K. Patricia Cross, discusses three strands of interest in assessment aimed at educational improvement. These three movements are: (1) state-level accountability; (2) institutional assessment; and (3) classroom assessment. The assessment movement appears to be approaching a confluence of these three streams. The second paper, "The Truth May Make You Free, But the Test May Keep You Imprisoned: Toward Assessment Worthy of the Liberal Arts" by Grant Wiggins, explores the shifts that would have to occur to make assessment really appropriate for a liberal arts education. Ten principles are offered to avoid the pitfalls of treating assessment in higher education as mere certification that a student possesses sanctioned knowledge. The third paper, "Assessment and the Way We Work" by Pat Hutchings, considers the relationship between assessment and learning in the real world of higher education. The real issue in education is not mere assessment, but rather the ways in which faculty work for learning. (SLD)

THE AAHE
ASSESSMENT
FORUM

AMERICAN ASSOCIATION FOR HIGHER EDUCATION

ASSESSMENT 1990:

UNDERSTANDING THE IMPLICATIONS

Papers By:

K. Patricia Cross
Grant Wiggins
and
Pat Hutchings

ASSESSMENT 1990:

UNDERSTANDING THE IMPLICATIONS

Papers By:

K. Patricia Cross
Grant Wiggins
and
Pat Hutchings

Presented at

The Fifth AAHE Conference on Assessment
in Higher Education
June 27 – 30, 1990
Washington, DC

Table of Contents

Streams of Thought About Assessment

by
K. Patricia Cross

Streams of Thought About Assessment*

by K. Patricia Cross
Elizabeth and Edward Conner Professor of Higher Education
University of California, Berkeley

The invitation to open this Fifth AAHE Conference on Assessment has provided me with the opportunity to think about both the origins and the destination of the assessment movement. Looking upstream, toward the origins of assessment, I see three rather different streams of thought. Looking downstream, I see these streams flowing together toward their ultimate destination of improved learning opportunities for students. Coming from California, which is experiencing its fourth consecutive year of drought, I am attracted to the imagery of full and bountiful streams which flow together at some point to make a powerful river. And it is that river and its tributaries that I want to explore tonight.

One doesn't have to look too closely to observe that the streambed of assessment has not always been smooth. Perhaps we should expect to find rocks and whirlpools, which in some places cause ripples in an otherwise serene higher education, and in other places cause turbulent white water as higher education wades thigh-high into assessment. Ultimately, of course, as we reach the confluence, we expect to be in the full swim of assessment and enjoying the challenge, having conquered not only our fear of drowning in data, but also our concern about floating aimlessly or—perhaps more frightening—being swept over the waterfall by forces beyond our control.

Before I get carried away in the watery drift of my own metaphor, let me try to find a viewing position on the bank where I can look upstream at the headwaters of these three rather different

*Presented at the Fifth AAHE Conference on Assessment, in Washington, D.C., June 27, 1990.

streams—all, however, flowing toward the same destination: the improvement of education for students.

First, there is the state-level accountability stream. The boaters in this stream are primarily governors, state legislators, and college and university trustees. Their most visible and immediate task is to fulfill their responsibility as guardians of the public trust in education. They must ensure that education is accountable to those who depend on and fund education—taxpayers, students, parents, employers, and in general all of us who make up a society which is increasingly dependent on educated citizens and workers for both its humanity and its productivity.

Assessment as accountability is illustrated in the remarks of Missouri governor John Ashcroft, who chaired the Governors' Task Force on College Quality. In 1986 he said, "The public has a right to know what it is getting for its expenditure of tax resources; the public has a right to know and understand the quality of undergraduate education that young people receive from publicly funded colleges and universities" (National Governors' Association, 1986, p. 154).

Since Governor Ashcroft made those remarks, the state-level accountability stream has quickened, but it has also changed direction; it now runs more toward improvement than accountability. It seems that the public is not so much interested in data that will document inadequacies as in action that will improve education.

So the turbulence once present in statewide assessment has calmed, and this vigorous stream with its headwaters in state legislatures now flows toward, if not actually into, the second stream—institutional assessment. Institutional assessment has many tributaries: it collects information from a variety of sources about academic achievement as well as student retention, graduation rates, and job placement. From this feedback, faculty and administrators hope to learn more about the impact of the total college experience on students.

The rationale for institutional assessment is illustrated in the remarks of John Richardson, president of DePaul University, who was quoted in a recent issue of the *Chronicle of Higher Education*. He said, "With tuition rates spiralling, parents and students are asking what we are contributing as Catholic colleges and universities. We have an obligation to tell them what we stand for and how we implement what we stand for" (*Chronicle*, May 23, 1990, p. A16).

In those words, President Richardson brings together several thoughts about assessment. First, accountability means not just fiscal accountability to taxpayers but a broader accountability, on the part of public and independent institutions alike, to students and to society. Second, President Richardson emphasizes the importance of institutional mission. The early state-level search for standardized tests or other common measures of assessment has moderated, in recognition that colleges have different missions and the measures used must be appropriate to the mission of the individual institution.

The third stream of assessment, classroom assessment, has its headwaters in thousands of classrooms across the country. The Classroom Research Project, located at the University of California,

Berkeley, and jointly funded by the Ford Foundation and the Pew Charitable Trusts, is a grass roots movement in which college teachers design the assessment measures themselves and engage in the continuous collection of information about student learning in their own classrooms (Cross and Angelo, 1988). The task of the classroom assessor is to continuously monitor student learning and to experiment with teaching techniques and strategies, searching for those that will improve students' learning. While the concepts and principles of classroom assessment are also applicable to administrators, counselors, librarians, and others who have direct responsibility for student learning in out-of-class activities, the Classroom Research Project has been working primarily with faculty development specialists and discipline-oriented classroom teachers.

These then are three major streams in the assessment movement. In my opinion, all three are necessary. They complement each other in their strengths and weaknesses, and fortunately, boaters on the three streams are beginning to shout across the banks to communicate with one another. I hope to end up this evening at the confluence of these streams of assessment activities, but I want to look first at the pollutants, crosscurrents, and fresh springs in each of the streambeds of the assessment movement.

The turbulent statewide accountability stream of assessment that educators once feared would overflow its banks has slowed as it becomes more mature. But it has also cut a deeper channel with a clearer sense of direction. Peter Ewell and his colleagues (1990), looking at state-level assessment initiatives, observe that "a much sharper image has emerged among state leaders" (p. 3). Whereas program review, admissions standards, retention, assessment of entry-level skills, and similar activities were once viewed under the broad umbrella of assessment, the focus has zeroed in on *learning* as the target. Most states now view assessment as part of a larger educational reform effort, and they are beginning to encourage institutions to design their own assessments, consistent with their own missions.

While direct involvement of state officers in the design of assessment measures and methods appears to have slowed in recent years, it is clear that the state stream is in no danger of being diverted or of drying up. Ewell and his colleagues conclude that ". . . state-level resolve is becoming both insistent and permanent" (p. 5). If institutional assessment appears to lose momentum, it is a near certainty—metaphorically if not geologically—that new waters will rush in from the statewide stream.

The statewide accountability stream makes an important contribution by insisting that colleges define their goals and design an assessment program to determine how well they are meeting them. Indeed, there is some evidence that the institutional stream is fed in part by the statewide stream. More than half of the institutions now engaged in assessment admit that their primary motivation is an existing or anticipated state mandate (El-Khawas, 1989).

It is sad but apparently true that in states where there is no legislative pressure for accountability, colleges seem less inclined to be innovative and progressive in this assessment movement. Linda-

Darling-Hammond observed at this Assessment Conference in 1988 that one of the few things on which social scientists agree is that "people will do more of whatever they are evaluated on doing" (1988, p. 1).

That finding applies to teachers and students as well as to administrators, of course. Given sufficient incentive, teachers will teach to the test, students will study for the test, and test scores will rise. There is a problem, however. Scholars from the National Academy of Education put it this way, "Unfortunately we are apt to measure what we can, and eventually we come to value what is measured over what is left unmeasured" (Committee of the National Academy of Education, 1987, p. 51). This argues that we should be very careful in our choice of accountability measures. Research is consistent in demonstrating that as scores on the instrument being used for assessment rise, scores on other measures tend to decline (Darling-Hammond, 1988). In the end, what is measured will change what is done. If we can measure critical thinking, for example, but not compassion or respect for people from different cultures, then we will teach to the test of critical thinking, and we will come to value critical thinking over other characteristics that are more difficult to measure and more difficult to teach.

There is already a suggestion that college faculty members tend to value what can be measured. This spring, my colleague Tom Angelo and I administered the Teaching Goals Inventory (TGI) to 2,800 faculty members in 33 two- and four-year colleges. The TGI asks teachers to select one course that they currently teach and rate the importance of each of 52 goals to the teaching of that course. Among the top-rated goals was the teaching of facts. A majority (52%) of all faculty members rated "learning the terms and facts of this subject" an "essential" goal of their teaching. Terms and facts are, of course, easy to measure. On the brighter side, however, 52% of the teachers also rated the development of analytic skills an essential teaching goal. Analytic thinking is somewhat more difficult to measure than mastery of definitions and facts, but the science of measurement has become much more sophisticated over the past decade, and most people would probably agree that we can measure analytic skills today with reasonable confidence.

For contrast, let us look at some of the less popular teaching goals. Fewer than one-fifth of the faculty respondents rated the following goals essential to their teaching: development of the capacity to make informed ethical choices, development of commitment to one's own values, and development of an informed appreciation of other cultures.

While I don't believe that this means that college teachers think these values are unimportant, it can be argued, I suppose, that such values are not well-addressed through academic subject matter, and thus are not the responsibility of college teachers. Some would go so far as to contend that since these goals are not easily measured, they cannot be taught with confidence in this assessment-conscious age. I would argue, however, that there are some things, not easily measured, that can be taught with confidence. For example, although only 20% of the respondents to the TGI said that "improving the ability of students to work productively with others" was an

essential goal for them, that skill would appear to be fairly easy to teach through collaborative learning—and collaborative learning is turning out to be an effective way to teach content, as well.

Finally, there are some things that we must try to teach regardless of whether or not we can measure them at present. The goal of cultivating an "active commitment to honesty," for example, was considered an essential teaching goal by fewer than one-third of our faculty respondents. What is more alarming, perhaps, is that 18% of the teachers said that cultivating a commitment to honesty was unimportant or not applicable to their teaching. (For the sake of comparison, fewer than 3% found the teaching of facts unimportant or irrelevant.) Ethics fared even worse than honesty; 19% considered helping students "develop the capacity to make informed ethical choices" essential, while 32% considered it unimportant or not applicable to their teaching.

Respect, ethics, and honesty appear to be endangered species in higher education. Items on the TGI having to do with these values were consistently selected as "essential" by only small minorities of college teachers. If we wait until we can measure them, however, they may disappear entirely. If such values are learned, as of course they are, where should young people learn them? Who should be held accountable for teaching them? Do we have any business trying to teach them if we can't measure whether or not students learn them?

So there is a crosscurrent in the accountability stream of assessment that presents dangerous navigating for the naive and unwary. To the extent that assessment-for-accountability is successful, education will become what we are capable of measuring. For the moment, at least, we applaud advances in measurement, but before we perfect our measuring devices, let us be sure that what we are measuring truly reflects what we want students to learn.

Certainly we want students to be able to analyze arguments, to think clearly, and, in general, to be able to apply their knowledge to the solution of a wide variety of problems. At the same time, we also want them to be concerned, compassionate human beings who use their knowledge to ethical and moral ends. The first set of characteristics is hard enough to measure; we need to work even harder on measuring the softer values of education.

Now let me return to the relative safety of the bank and look upstream at institutional assessment. The fresh spring in the streambed of institutional assessment is the notion that student learning can and should be monitored, and that educators have an obligation to inform themselves of how well they are doing, whatever they purport to be doing. Institutional assessment also has the potential for a powerful flow of shared values when members of the campus community join together in conversations about institutional goals and how effectively teaching and learning are meeting them.

Where institutional assessment runs across a dry creek bed is in turning assessment data, reports, and recommendations into a full and bountiful flow of action. Even if the data are good, the recommendations sound and do-able, and the reports widely distributed to the appropriate people—which is much more

than most institutional assessments accomplish at present—there is still a considerable gap between the collection of data and the improvement of education. In fact it becomes quite difficult to diagnose problems and to pinpoint responsibility for addressing them.

Let me illustrate the problem with a typical scenario: In most institutional assessments, it is assumed that if the faculty, as a group, are made aware that students lack knowledge in subject matter areas considered important, the collective faculty will take steps to correct the deficiency—usually through *adding* of courses or requirements. This leaves the impression that education is additive; that is, that more requirements equal more learning. Virtually ignored is the premise that much of what is taught is not learned. Yet learning, correctly understood, is transformational rather than merely additive; it consists of new learning transforming what already exists in the minds of learners to lead to deeper understandings and appreciations. The feedback from institutional assessment is usually interpreted as a problem in curriculum; it rarely addresses problems in instruction.

Thus, there are two rather large rocks which constitute obstructions as the stream of institutional assessment moves toward educational improvement. One is the problem of deciding what to do; the other is pinpointing who should do it. If the assessment shows that students are poor in critical thinking, for example, does that mean that they need more courses that emphasize critical thinking, or that teachers of all subjects need to be more aware of teaching to the goal of critical thinking? Does the problem lie in the curriculum, or in the effectiveness of instruction?

Although projects emphasizing "critical thinking across the curriculum" hold that all teachers should be concerned about teaching analytic skills, our data from the TGI show marked and consistent differences in what teachers from the various disciplines say they are trying to accomplish through their teaching. For example, whereas 73% of the math teachers say that the development of analytic skills is an essential goal of their teaching, only 43% of the social scientists assign it such high priority. In general, teachers in the humanities and social sciences give much higher priority than teachers of mathematics and science to a cluster of TGI items that we might label "General Education Values." In that cluster are items such as these: development of a lifelong love of learning, openness to new ideas, informed appreciation of other cultures, capacity to make informed ethical choices, and so forth.

If institutional assessment shows that students score low on some measure of general education values, does that mean that students should take more courses in the humanities and social sciences and fewer in math and science? Or does it mean that math and science teachers should be helped to emphasize humanistic values in the uses of science? Or does it mean that a special core of teachers should develop courses in general education to teach the desired values?

Take another example. Let us assume for the moment that institutional assessment shows that students are not adequately prepared to think creatively. There are several ways to give students

more practice in creative thinking. One is to require more courses in subjects taught by teachers of English and the visual and performing arts who, according to our TGI data, assign a high priority to creative thinking. Another is to institute a faculty development program aimed at helping all teachers to teach creativity more effectively. The question is whether some disciplines teach creativity more effectively than others (and whether creativity is a general trait or discipline-specific) or whether some teachers teach it more effectively—or perhaps it is a combination of factors arising from the observation that people who value creativity are likely to enter disciplines that require or encourage it.

Now let me look at the third stream of assessment. Classroom assessment, as originally conceived (it changes a bit as it reaches the confluence), involves individual classroom teachers in the assessment of student learning in their own classrooms. In contrast to the other two streams of assessment, the emphasis is on the improvement of classroom teaching more than on accountability or curriculum. The assumption is that if teachers find out that students are not learning what they think it is important to teach, then teachers will experiment with a variety of teaching techniques to improve the performance of their students.

The fresh springs in this assessment streambed are several: First, confidence in the ability of discipline-oriented teachers (as opposed to measurement experts) to design appropriate assessment measures. Second, confidence that college faculty members will use the results of their assessments to experiment with and modify their teaching. Third, the possibility that a critical mass of individual teachers, carefully observing learning and concerned about how they can improve it, will change the institutional climate as well as the learning environment in classrooms across the campus.

Thus classroom assessment ensures that teaching faculty are personally and intellectually involved in assessment; that they receive feedback which they perceive as directly relevant to their teaching goals; and that such feedback has direct and immediate implications for their teaching.

Our experience with classroom assessment to date suggests that as individual faculty members become involved, conversations about teaching increase, curiosity about learning quickens, and new energy is pumped into the environment. There is contagious enthusiasm in the feedback we are getting from classroom assessors. One writes, "Now I couldn't teach without it; I depend on the feedback. What I especially like is how classroom assessment has the character to change gradually the whole teacher/student interaction, reducing distance, encouraging student participation and the assumption of responsibility for their own learning." Another says, "It is simply a part of what I do now."

Now I am ready to look downstream. Do we have the possibility of a clear and bountiful river at the confluence of the separate assessment streams?

I want to look first at the merging of classroom assessment and institutional assessment. Institutional assessment is more than the aggregation of hundreds of separate classroom assessments, and institutional improvement is more than

the sum of individual modifications in teaching. While involving everyone in classroom assessment is a strong step toward improvement, it is probably not sufficient to maximize institutional effectiveness. At the same time, institutional assessment is lifeless and often stagnant without the personal involvement of large numbers of faculty members.

What do we get if we merge the two powerful streams of classroom assessment and institutional assessment? One possible outcome is collaborative classroom assessment. One of the most interesting aspects of our experience with classroom assessment is that it appears that teachers like to use the results of their classroom assessments to initiate—often for the first time in their teaching careers—conversations with their peers about teaching and learning.

I confess that I am mildly surprised by the strength of this tendency toward faculty collaboration in classroom assessment. One of my original motives in proposing classroom assessment was to provide individual teachers with a course of action that fit well into the traditions of academic freedom and the presumed privacy of the classroom. Here at least was something that teachers could do on their own without going through committees, or applying for a grant, or seeking anyone's blessing. It appears, however, that the presumed sacred tradition of privacy in the classroom is more myth than desire. Teachers are eager to collaborate on the design of classroom assessments and to compare results and share interpretations.

The confluence of classroom assessment and institutional assessment seems natural and inevitable. There may well be a cumulative effect of classroom assessment that is more than the sum of the parts. If classroom assessment did nothing more than cause individual teachers to make even modest improvements in their own teaching, the aggregated effect on students—with eight teachers per year, each monitoring learning carefully and striving to improve it—would be considerable. Collaborative classroom assessment, however, expands the circle of participation from teachers and students in single classrooms to teachers talking with teachers and perhaps even students talking with fellow students about learning assessments across classrooms and disciplines.

It is clear to us now that three kinds of collaboration appear to be especially productive. One is based on common disciplinary interests, another on shared values, and a third on stimulating differences.

Departmental or discipline-based assessments are growing in popularity, but it is my impression that there are two rather different emphases, both valid—but different. One is the attempt to determine the extent to which the department contributes to the educational goals of the institution; the other is to assess how well the department achieves the goals of the discipline. These two emphases are not very different from earlier statewide discussions over whether to hold institutions to a set of common goals or to acknowledge diverse institutional missions.

Let me try to illustrate the use of collaborative classroom assessments within the department. The first step might be clarification of the goals of the department through administering something

like the Teaching Goals Inventory to departmental faculty. I can see some very spirited discussions as individual faculty members compare their teaching goals with those of their peers in the discipline. A next step might be the collaborative design of assessment devices to measure the accomplishment of high priority goals for the department. Individual faculty might then determine how close they are to the goals of the department by administering the assessment devices in their own classrooms. While there are admittedly pressures to conform to the group goals, we should keep in mind that the individual goals of a mildly eccentric professor may be quite valuable to students and to the institution. Therefore, although our early experience with classroom assessment shows more eagerness to share data than to keep it confidential, I recommend that departments and institutions refrain from identifying the data from any individual classroom assessment. It is probably sufficient that individual teachers have thought about their own teaching goals, have collected data showing their departure from departmental norms, and have made conscious decisions to conform, or not to, in goals as well as in teaching strategies to accomplish those goals.

Let me make this example more concrete by illustrating what might happen in our TGI prototype of an English department. We had 327 English teachers in our TGI sample this spring, constituting 12% of our total sample. There is a cluster of 5 items, out of the 52 on the TGI, which 60% or more of the English teachers consider "essential" to their teaching. In order of importance, they are: improve writing skills; develop capacity to think for oneself; develop analytic skills; develop ability to think holistically, i.e. to see the whole as well as the parts; and develop ability to synthesize and integrate information and ideas. Now let us look more closely at just the top three items which were considered "essential" by two-thirds or more of the teachers of English.

- Improve writing skills: 84% for English teachers—significantly higher than the 61% of all faculty respondents who considered writing their top priority.

- Develop the capacity to think for oneself: 75% for English teachers; 55% overall.

- Develop analytic skills: 66% for English; 53% overall.

With such a high level of departmental agreement, it is appropriate to ask how English faculty might assess whether they are accomplishing their goal of developing the capacity of students to think for themselves. A tough one? Not necessarily. It would seem appropriate to ask students to write their own interpretation of the balcony scene in *Romeo and Juliet,* for example—how that interpretation might differ between men and women, and how their interpretation might be related to their own personal/cultural background or experience. In that exercise, students tackle writing, thinking for themselves, and analysis from several different perspectives. Does this exercise constitute assessment? I think so, and even more important, it is teaching in the interests of improving student learning on goals that are considered important.

But what about the 15% of the English teachers in our sample who departed from these typical English Department norms and gave "thinking for oneself" relatively low priority? What do they teach? Are they skills-oriented teachers—grammar, spelling, or ESL, perhaps? Should they be teaching for broader goals? Seeing how far they depart from the norms of English teachers should raise questions in their own minds, but these teachers may legitimately conclude that their goals are the right ones for them and for their students and their subject matter.

There is a second way in which classroom assessment can merge with institutional assessment: Faculty may cluster around shared values or goals that do not follow disciplinary lines. Much of this sort of merging is already taking place. Writing across the curriculum, critical thinking, general education, and more recently the teaching of ethical decision making, all attract clusters of faculty with common teaching goals.

Teachers with common teaching goals might form support and study groups, taking on the task of devising appropriate assessment measures and administering them in their own classrooms to determine their effectiveness in teaching toward goals that they value. How does one teach writing or critical thinking or a lifelong love of learning? How do students learn these things? Experimenting with teaching techniques that seem especially effective and sharing teaching experiences and insights into how students learn should prove a useful and effective approach to professional development.

The third faculty cluster was called to my attention recently by a faculty member who was part of a faculty-development effort to train college teachers in the methods and philosophy of classroom research. This group of about 20 teachers found common ground, not necessarily in their discipline or in their teaching priorities, but in their common interest in classroom assessment. They were remarkably creative in devising classroom assessment techniques, and they found their differing experiences and disciplines mutually stimulating. It was the first opportunity many of these teachers had had in their entire teaching careers to talk in some depth with teachers from other disciplines about their priorities and goals as teachers and about how they assessed their own effectiveness. Their observations about how students learn were supported and enriched by the observations of teachers looking at learning outcomes from different perspectives.

I am sure there are more opportunities for collaborative classroom assessment than I have described here. Indeed, the more dyads and triads and informal and formal groups of all sizes and constituencies a college could encourage, the more stimulating and productive the environment would become. When all is said and done, the most far-reaching and long-lasting reforms of undergraduate education will come when individual faculty members or small groups of instructors adopt the view of themselves as reformers within their immediate sphere of influence, the classes they teach every day.

The picture I have tried to paint this evening shows three streams of assessment moving along in their own streambeds, each with a few diversionary rocks, some with whirlpools using a lot of

energy to make very little progress, and some with shoals at various points along the route to improved education for students. At the confluence of these streams, there is a deep and powerful river. I think we are approaching that confluence in the assessment movement now.

In closing, I have just a few words of advice. To those who want to look good in this assessment movement while sitting on the beach, look out, you may get burned unless you are fully covered. To those who would like to look like you are in the swim of assessment without putting forth too much energy, it might be wise to learn a few strokes. The crawl is serviceable, but the butterfly is more spectacular. To those of you who are paddling your own canoe pretty far upstream, you might wish to consider finding a partner for a two-person kyack or a group for a 10-person racing shell. For the very few power boaters in this crowd who like to churn up the waters, leaving a trail of data in your wake, don't look back. The rest of us may be gaining on you. ■

References

Cross, K. Patricia, and Angelo, Thomas A. *Classroom Assessment Techniques: A Handbook for Faculty.* Ann Arbor: National Center for Research on the Improvement of Postsecondary Teaching and Learning (NCRIPTAL), University of Michigan, 1988

Darling-Hammond, Linda. "Assessment and Incentives: The Medium Is the Message." AAHE Conference on Assessment, Chicago, Ill., June 8-11, 1988.

El-Khawas, Elaine. *1989 Campus Trends Survey.* Washington, D.C.: American Council on Education, 1989.

Ewell, Peter; Finney, Joni; and Lenth, Charles. "Filling in the Mosaic: The Emerging Pattern of State-Based Assessment." *AAHE Bulletin*, April 1990.

National Academy of Education. A Review of *The Nation's Report Card.* Cambridge, Mass.: Harvard University, 1987

National Governors' Association. *Time for Results.* Washington, D.C., 1986.

The Truth May Make You Free, But The Test May Keep You Imprisoned: Toward Assessment Worthy of the Liberal Arts

by
Grant Wiggins

The Truth May Make You Free, But the Test May Keep You Imprisoned: Toward Assessment Worthy of the Liberal Arts*

by Grant Wiggins
Director of Research
Consultants on Learning, Assessment, and School Structure

I confess that I have been kicking myself for getting involved with this topic. The more I looked at the title I obligated myself to, the more nervous I became. For one thing, I do not know as much as I should about current efforts to assess higher education. For another, I do not think anybody in his or her right mind can address this topic intelligently in so little time. And third, I think that the problem specified in the title of my talk is an insoluble one. It confronts us with one of many inescapable dilemmas about the liberal arts: the freedom of thought to go where it will versus the apparent need for uniformity in the testing process. In other words, we are not going to "solve" the assessment problem in the liberal arts, now or later. We are going to negotiate it—painfully; we are going to have to deal with some frequent, uneasy compromises.

So what I intend to do is a bit more modest than perhaps it first seemed. My aim is not so much to lay out a complete vision, but to give you my sense of the subtle but profound shifts that would be required if we were going to be serious about assessing for a liberal arts education. Second, I am going to propose to you a set of principles that we might call upon when liberal education is jeopardized—as it always will be—by an overly

*Presented at the Fifth AAHE Conference on Assessment, in Washington, D.C., June 29, 1990.

utilitarian or vocational view of teaching and learning. I think of these principles as mere first cuts, but perhaps they can hold you in good stead on a rainy day. Third, I will offer what I hope are some provocative and useful illustrations of alternative forms of assessment that befit the liberal arts. Many of the examples happen to come from the K-12 arena but nonetheless apply to your situations as well. I would encourage you, therefore, to resist a common nasty, little habit. If I should make reference to a fifth grade teacher's example, try not to be snooty about it. It is harder than you think to resist the feeling, and harder still to develop the almost anthropological mindset that enables one to find insight into one's own teaching from very different places in the system.

Let us begin thinking about dilemmas in education by returning to the first known assessor in the liberal arts. I am thinking, of course, of Socrates, the Socrates of the dialogues of Plato, where we regularly see those who either appear to be or profess to be competent put to the "test" of question, answer, and—especially—sustained and engaged conversation. (The dialogues themselves, of course, are filled with dilemmas. Many of them are left unresolved: a reminder of how these arts are meant to lead to questions, not answers—the little burrs that get under your saddle.) Socrates the assessor: he is certainly a strange one. He does not seem to have nice answer keys or scoring rubrics by his side. Yet I think that there is something to learn from thinking about assessment from a Socratic point of view.

I would like to view these issues through my favorite piece of literature, the dialogue called "Meno." Some of you no doubt know it. Meno, a brash young fellow, comes up to Socrates. The first line of the dialogue is, "Tell me, Socrates, how do we become virtuous?" In other words, he is asking how morality develops: through upbringing? moral education? by nature? Socrates responds in a very annoying and typically Socratic way. He says, "Well, I cannot answer that question. I do not even know what virtue is." Meno is clearly astonished to think that this could be possible, that a *bona fide,* certified sage does not know what everybody knows, namely, what it means to be good. But of course, after Meno makes the foolish mistake of venturing to tell Socrates what virtue is, Socrates proceeds to undress him two or three times.

Finally, in exasperation, Meno says a terribly revealing thing that goes to the heart of the distinction between conventional assessment done well and an assessment for the liberal arts. Meno says, "Well now, my dear Socrates, you are just what I have always heard before I met you. Always puzzled yourself and puzzling everyone else. And you seem to me to be a regular wizard. You bewitch me. You drown me in puzzles. Really and truly my soul is numb. My mouth is numb. And what to answer you I do not know." And here is the important part. "Yet I have a thousand times made long speeches about virtue before many a large audience. And good speeches, too, as I thought. But I have not a word to say at all as to what it is."

Meno's comment (and indeed the progress of the whole dialogue) ironically reveals what so differentiates conventional academic mastery from excellence befitting the liberal artist. Meno is

reduced to speechlessness, he thinks, because of the sophistry of Socrates's questions and analyses; the thoughtful reader knows, however, that Meno does not know what he is talking about. And yet Meno is a conventionally successful student. How do we know? Throughout the dialogue Meno is constantly dropping references—the ancient equivalent of footnotes—to all the famous people who say this and that about virtue, which he, of course, agrees with. And it is no doubt the case that Meno could be a successful speaker—effective, convincing. The point of the dialogue, of course, is that such rhetorical skill using borrowed ideas is not understanding; competent presentations are not sufficient. That is not what a liberal education is about.

What Socrates wants us to see—what Plato wants us to see by the way in which the dialogue is written—is that the conventional view of education is actually quite dangerous. If one gets better and better at what one does, one is less and less likely to question what one knows. Meno has been a dutiful student. (We are also meant to know that his name is a pun: It is very close in Greek to the word for memory: menon - mnemon.) Meno is an effective memorizer, able to make effective speeches with references to famous people. Isn't that what too much of our assessment is already about? Don't we too often fail to assess whether the student can do anything more than borrow quotes, facts, and figures?

But we also know from history that the real Meno was a nasty fellow: clever, ruthless. We are meant to know that. Because there is ultimately a lesson to be learned about "control" over knowledge and the ends to which "mastery" is put. Liberal education can never co-exist happily with other, more "practical" views of education because a liberal education is about rooting out thoughtlessness—moral as well as intellectual thoughtlessness.

There is, alas, such a thing as "thoughtless mastery" (as I have elsewhere termed it) and our syllabi and assessments tend unwittingly to reinforce it. Many of our students are quite good at this thoughtless mastery; you all know it. You know those looks in class, those mouth-half-open looks, the eyes slightly glazed; when people are fairly attentive but the brain does not seem to be quite engaged; when, alas, their eyes only focus to check scores on other people's papers, and to press you for extra points here and there.

Paradoxically, many professions *require* unthinking mastery—and run the risk of an amoral technical approach to life. I think we forget this. I do not want the pilot who flew me to Washington to be questioning his knowledge or his existence. Nor do I want my brain surgeon to be thinking about what virtue is. One of my passions is baseball, and I was recently reading George Will's new book called *Men At Work* on the craft of playing and managing major-league baseball. There is an odd but insightful phrase in it about this kind of thoughtless mastery that rings quite true. The good hitters talk about not thinking too much—that it is very dangerous to do so. Rather, what has to take over the hitter is something called "muscle memory"—a wonderful phrase for the kind of unthinking skill that we admire in athletes.

There is no reason, however, for colleges and universities to assume that their job is to promote unthinking mastery of others' ideas (while also abetting the

other forms of thoughtlessness that too easily follow). Colleges are derelict, I think, in giving up the only sanctioned time when we have *a moral obligation to disturb students intellectually.* It is too easy nowadays, I think, to come to college and leave one's prejudices and deeper habits of mind and assumptions unexamined—and be left with the impression that assessment is merely another form of jumping through hoops or licensure in a technical trade.

Certainly we *say* we would like to see more "real" thinkers, and we bemoan doltish behavior in our students. I think we protest too much. Our testing and grading habits give us away. If you do not believe me, look how often students give us back precisely what we said or they read. On the other hand, you should not think that I mean rigor does not matter. That is part of the dilemma. The great mistake that has been made in school reform by many so-called progressives, and by much of the alternative schools movement, is to assume that to be liberated is to be liberated from discipline. That is a mistake, and it is one reason why alternative school people end up shooting themselves in the foot: because they produce a lot of free spirits who are not always very capable. If I had to choose, I might go with the alternative schools, but it is a bad choice and it shows that we have not negotiated the dilemma in K-12.

So we have to think about rigor. We have to think about alternative assessment as more than just engaging students better, which it invariably does (you know this if you have done simulations, case studies, portfolios, or dramatic presentations with your students). We need more than engaging activities. We need truly standard-setting and standard-revealing assessments. Or as psychologist Lauren Resnick puts it: What we assess is what we value. We get what we assess, and if we don't assess it, we won't get it. True about rigor, but also true about the intellectual virtues.

Some of you know, if you have read some of the things that I have written on alternative assessment, that one of my definitions of authentic assessment is that it is "composed of tasks that we value." It is not a proxy. It is not an efficient system to shake out a grade. Efficiency and merely technical validity as the aims of assessment will *undermine* liberal education. Rather, the test should reveal something not only about the student but about the tasks and virtues at the heart of the subject—its standards. But it is damn hard to design tasks to meet those criteria. It is very easy to score for efficiency and to look at what is easy to score rather than what is essential.

Let me cite three other dilemmas before giving you some principles and examples of how we might think about assessment that would do justice to the liberal arts. The first dilemma, confronting you more often as a teacher the higher up you get in the system, is whether to stress students' mastery of the ideas of others or mastery over their own emerging ideas. In fact, we do believe that it is important for students first to control subject matter and acquire skill within the discipline before they get "creative." Or to paraphrase Thomas Kuhn, one must have complete control over the existing "paradigm" if dramatic new paradigms or original thoughts are to occur.

Whatever Kuhn's merits as a historian

and philosopher of science, I think he is dead wrong about education. I think it is terribly important that would-be liberal artists immerse themselves, from the word go, in questioning the paradigm *as* they learn it: They should study it, poke it, prod it, and not wait until they have mastered it—because you can have a *long* wait. And many of your bright and able minds are likely to drop out mentally or physically because they cannot wait that long. Conversely, the ones that stick around may be more dutiful than thoughtful.

Inevitably, if we first demand control over the subject matter in its entirety, we run a moral as well as an intellectual risk. We run the risk of letting the student believe that Authority and authoritative answers matter more than inquiry. We may well end up convincing students that "Knowledge" is somehow something other than the result of personal inquiries built upon questions like theirs. And in fact, many students do believe that: There is "Knowledge" over there and there are "questions" over here and never the twain shall meet.

A second way to put the dilemma is more classic: useful versus useless knowledge. There is an important sense in which the liberal arts *are* useless, summed up in that little comment supposedly made by Euclid 2,000 years ago when someone complained that geometry was not good for very much. He said, well, give him three drachmas if he has to get some usefulness out of studying it. But there is a more important truth in this desire. It is not at all clear that this unending inquisitiveness and poking over, under, and around knowledge is useful. Indeed, I can tell you from working with adolescents for so many years (prone to outbursts of honesty and not feeling the need to appear like eager apprentices), that many of them regard it as profoundly useless. On the other hand, we must ourselves keep clear the distinction between "useful" (or "relevant") and "meaningful." Students are *not* entitled to usefulness in a liberal education, but they *are* entitled to a meaningful encounter with essential ideas. We often disappoint—either by pursuing ideas that are *too* relevant but transitory; or by being insensitive to their need for provocations, not packages of predigested "knowledge," to chew on.

Third, we have to recognize that the urge to shun the liberal arts may have a great deal to do with the essential urge to feel competent. People go to school, it seems to me, indirectly to feel good about themselves. They want to develop competence because they want to develop confidence—or is it the reverse? The trouble with a liberal education is that it does not satisfy that need at all. It is unpleasant. It is disturbing. Many people drop out mentally and become hypercompetent because they cannot deal with the ambiguity and uncertainty that is the hallmark of the liberal arts.

Well, then, suppose I am right about this. Suppose we are in danger of treating assessment in higher education—as we are now increasingly in danger of treating assessment in lower education—as certification that a student possesses sanctioned knowledge. Where would we look for effective alternative strategies? How can we highlight the liberal arts side of the dilemma? What principles might guide us in designing assessments for the liberal-artist-in-training?

Let me offer you 10 tentative principles.

Principle #1. The heart of the liberal enterprise is not a mastery of orthodoxy but learning to justify one's opinions. Because the modern university has its roots in the Middle Ages and in religious training, it is built upon an irresolvable tension between orthodoxy and the promotion of inquiry. We tend to forget that. To this day, it seems to me, we still lean pretty heavily on the orthodoxy side: Up until the graduate experience, students have first to demonstrate their control over other people's knowledge. Yet we would be wise to begin our reforms from the perspective of the ultimate educational experience with which we are all familiar: the dissertation and oral in defense of a thesis. We should think of all assessment as designed primarily to give students an opportunity to *justify* opinions that are being developed as they explore subject matter.

This implies that one of the most important things that we can do in assessment is to examine the students' response to our follow-up questions and probes of their ideas. It implies, for instance, in assigning a paper and evaluating it, that the student should have to formulate a response, to which we then respond as part of the formal assessment process, not as a voluntary exercise after the test is over or the paper done.

Taken to the limit, I would argue that one of the most important things that we can do with students is to assess them on their ability to punch holes in our own presentations. They have a right to demand from us justification of our point of view. That is what the liberal spirit is about. It sends a moral message that we are both, student and teacher, subservient to rational principle.

Principle #2. The second principle is that we really need to think of the student as an apprentice to the liberal arts. And like all apprentices, students should be required to recognize and produce quality work. They should not get out of our clutches until they have produced some genuinely high quality work. Now, what do I mean by that? Well, it is really a subtle shift in thinking. We all expect quality as teachers, but I do not believe that we demand it.

For instance—and here is one of those sixth grade examples—there is a teacher in Louisville who in one of her first assignments to her social studies students demands that every student read a book and do a book report. Not a particularly interesting task, but what is fascinating is what she demands. She demands that the paper be perfect. She demands that the students not turn it in until it is. She demands that they seek out anyone and everyone who will help them make it perfect.

Well, needless to say, the kids freak out. Especially the bad ones who are convinced that they cannot produce quality work. To make a long story short, they do. Oh, we could quibble with the idea of a perfect paper, but the kids understand full well what is meant. They really do. It is quite something to see. They understand that they have to ratchet up the seriousness with which they work. That they cannot wait to find out the quality of the work they produced. That they have to produce the quality work *first.* Making a point that many of you know is now critical to the alternative

assessment conception: Assessment and self-assessment must be intertwined if we are serious about empowering people. To demand quality is also to structure assessment so that the student does not merely have the opportunity to rehearse, revise, rethink, but is actually required and expected to do so.

One of my favorite assignments when I taught at Brown was to ask students for their final paper to rewrite their first paper, based on all they had since learned or thought. A number of the upperclassmen told me that it was the most important event in their four years at Brown. They were astonished to see how their thinking had changed. They were astonished to discover how sloppy that early work seemed to them in retrospect. In short, they were learning about quality.

Further, they were learning about thinking, that thinking does not stand still and should not. Demanding quality, in other words—and this is part of the shift in thinking that is required—means we begin to focus our assessment on what Aristotle called the intellectual virtues. Does the student display craftsmanship, perseverance, tolerance of ambiguity, empathy when everyone else is critical, a critical stance when everyone else is empathetic? Can the student, without prodding, re-think and revise a paper or point of view? A liberal arts education is ultimately about those intellectual virtues. When all of the knowledge has faded away, when all of the cramming has been forgotten, if those intellectual dispositions do not remain, we have failed.

Now, some people get very squeamish about assessing things like perseverance, style, craftsmanship, love of precision. I do not. If we value it, we should assess it. That does not mean that we are arbitrary. That does not mean that we are subjective. Yes, we have to worry about validity and reliability. In fact, what I think it means to assess habits of mind is not to directly score them at all. But rather to devise tasks that *require* the habits we value.

My metaphor for this is "Outward Bound." Assessment should be like intellectual Outward Bound. It should reveal to the student what we value as traits in them by the virtues required to accomplish the task at hand. It should not be possible to do an end-run around those habits; students who can get *A*'s by missing class and cramming are telling you something about the failings of your assessment system.

Sometimes it is as subtle a shift as sending the message day in and day out that quality matters and you are held accountable for quality. One of my favorite little tricks in that regard comes from Uri Treisman at Berkeley and his work with minority mathematics students. He demands that every piece of work the students hand in be initialed by another student; students get a grade both for their own paper and for the paper on which they sign off. This sends a message loud and clear that quality matters, that you are personally responsible for quality, and that it is in your interest to find out about quality *before* hearing from the authority. Quality control is about avoiding poor performance before it happens.

Principle #3. This leads directly to principle #3, a point familiar to many of you who have been at this kind of work, but one that cannot be made often-

enough. A liberal arts assessment system has to be based on known, clear, public, non-arbitrary standards and criteria. There is no conceivable way for the student to be empowered and to become a masterful liberal artist if the criteria and standards are not known in advance. The student is kept fundamentally off balance, intellectually and morally, if the professor has a secret test and secret scoring criteria.

Consider the performance world, as opposed to the academic world, and how much easier it is for performers to be successful because of this very basic fact. The test is known from day one. The music, the script, the rules of debate, the rules of the game are known: Genuine mastery in the performance arena means internalizing public criteria and standards until they become one's own. Unfortunately, in education, especially in higher education, there is a vestige of our medieval past, when tests were a bit of mystery and novices had to divine things. I was disappointed to learn when I was a teaching assistant at Harvard that most undergraduates are still not allowed to see their blue books. And then somebody told me that at Oxford and Cambridge they burn them.

I think this is an unfortunate and deadly tradition. It is also a legacy of tests used as gatekeepers, not as equitable vehicles designed for displaying all that a student knows. Most people in this room, I suspect, would say it is the *student's* responsibility to figure things out, to respond to the test as the test demands, and to produce quality work on *our* terms. I am not convinced of that. Why isn't the university required to meet students halfway and give them a chance to reveal their strengths and play from their strengths? It would be as simple as giving people the option of alternative forms of doing the same assignment.

But I think it runs much deeper. We are still using testing as a sorting and categorizing system. And elitism should not be confused with meritocracy. Our most common habit in scoring and grading, namely scoring on a curve, is unjustifiable in my view. Its sole purpose is to exaggerate difference rather than reveal strength. It makes our life easier and it relieves us of justifying the grades and scores that we give. It is needlessly debilitating—as opposed to a challenge that we can rise to when we know, understand, and appreciate the criteria.

Of course, many of you know the solution. Scoring rubrics, model papers, videotaped model performances, anything that can give students an insight into, allow them to enter the field and acquire its standards, by seeing exemplary performance *before* they do their work. I do not know why in the world we keep such matters a secret. It is cuckoo—and dysfunctional.

Principle #4. It follows that what a liberal education is about—and what assessment must be about—is learning the standards of rational inquiry and knowledge production. And this implies that self-assessment is a critical and early part of assessment. Now, many of you know about Alverno's use of self-assessment and it has been borrowed by many of us. I just want to give you one of my favorite Alverno examples because I think it illustrates so well different ways of thinking about this.

One of Alverno's goals for students is

competency in oral communication. Early on, a student must give a videotaped talk and so one's first hunch is, oh, well, you are going to assess the talk. No. After the student gives the first talk and it is videotaped, the student is assessed on the accuracy of the self-assessment of the videotaped talk. That is a fundamental shift in point of view. If we want people to gain control of important habits of mind and standards, then they have to know first of all how to view those things accurately and apply criteria to their own work, and not always depend upon another person to do that.

It is also a basic lesson in habit development. You have to know what you are *supposed* to be doing before you can do it. And that knowledge is crucial in making you stick with it and believing that it is possible. Otherwise I do not think any of us would quit smoking or lose weight. It suggests as a practical corollary that no major piece of work should get turned in without some self-criticism attached to it. And that self-criticism should be assessed for *its* accuracy.

Principle #5. Most education, it seems to me, treats the student as a would-be "learned spectator" rather than a would-be "intellectual performer." The student must metaphorically sit in the bleachers while others, mostly professors and writers of textbooks, perform. This arrangement takes us back to the idea that competency involves just remembering and applying what others say. It has dangerous consequences because it induces intellectual passivity. In an education for a would-be performer the student would experience the same "tests" that face the expert in the field—having to find and clarify problems, conduct research, justify one's opinion in some public setting—all while using (other people's) knowledge in the service of one's own opinion.

Let me give you a couple of my favorite examples of this. One of the finest classes that I have ever observed at any level was at a high school in Portland, Maine, where a veteran teacher offered a Russian history course. The entire syllabus consisted of a series of chronological biographies. It was, however, the *student's* job to become each person, in turn, in two senses: through a 10-minute talk, and then a simulation. After four or five students had presented their talks (and been assessed by other students on their talks), they had a Steve Allen "Meeting of the Minds" press conference which was chaired by the teacher; the "journalists" were the other students. Each party scored the other for its performance.

Now, I do not know about you, but I have sat through a lot of dreary reports. These were not dreary. In fact, they were as interesting and informative as any reports I had ever heard. I went up to the teacher and said, golly, how did you get them to do that? He said, well, it was very simple. There were only two criteria by which they were going to be judged and they were (a) whether the talk was accurate, and (b) whether it was interesting. This was real performing and using knowledge.

Principle #6. This one follows from #5. A liberal artist, if he or she has "made it," is somebody who has a style. Somebody whose intellectual "voice" is natural and clearly theirs. Read the turgid prose that we receive and you know that we are

failing to develop style, voice, and point of view. (Read our own writing in journals. . . .) Students are convinced we want the party line, and that the quality and insight possible in compelling prose is not necessary. It is an option.

There would be a simple way to get at this. After writing a lengthy research paper with all the requisite footnotes and bibliographical information, the student could be asked to turn the paper into a one-page piece to be delivered, in an engaging and insightful way, to an audience of laypersons. But it is not just an aesthetic issue, this business of style. It is a question of one's inner voice. One's intuition. The seed of a new idea that is easily crushed if it is not allowed to be heard. All of these are related to the idea of conscience, and, of course, it is no coincidence that Socrates talked about his little voice.

It is easy for students in American universities to lose that little voice. But that little voice is not just a "personal" voice irrelevant to "academic" accomplishment. It is the voice of common sense. It is the voice that can turn around and question the importance of what one has just spent two months working on. It is the little voice that says, ahh, come on, is this really *that* important? Or it is the little voice that says, you know, there is probably another way to look at this. It is the little voice that says, I have a feeling that there is something behind what the professor is saying, but I do not know enough to really pursue it so I will not. It is the little voice that most of us do not hear in our students unless we ask for it. An assessment should ask for it.

Such assessing need not be difficult. I saw an English teacher do it. In using peer editing, he told his students that they should reject and turn back any paper that was boring or slap-dash—and mark the exact spot in the paper where they began to lose interest. Nothing sends a message faster to students about writing and its purpose and quality. Nothing sends a message quicker that technical compliance with criteria is not always of primary importance.

There is another point to be made about voice and style. The thing that is so ghastly about academic prose is that one really does sense that it is not meant for an audience. And, of course, sometimes it is not. It seems to me that if we are serious about empowering students, we must get them to worry about audience in a deeper way. We must demand that their work be *effective*. We must demand that it actually reach the audience and accomplish its intended purpose. There is nothing more foolish, in my view, than saying, "Write a persuasive essay" without the students having to persuade anybody of anything. So let us set up situations in which the student *has* to persuade readers, or at least get judged by an audience on more than just accuracy. Even Socrates knew, in the clash of Reason and Rhetoric, that teaching had to be not merely truthful but effective.

Principle #7. Too often in assessment we worry about whether students have learned what we taught. This is sensible, of course. But let me take an unorthodox position. Such a view of assessment, taken to extremes, is incompatible with the liberal arts. One important purpose of those "arts that would make us free" is to enable us to criticize sanctioned ideas, not merely re-tell what was taught.

A less confrontational way to make the point is to remind ourselves that it is the astute questioner, not the technically correct answerer, who symbolizes the liberal artist. The philosopher Gadamer (with an explicit homage to our friend Socrates) argued that it is the dominant opinion that threatens thinking, not ignorance. Ensuring that students have the capacity to keep questions alive in the face of peer pressure, conventional wisdom, and the habit of their own convictions is what the liberal arts must always be about.

Admittedly, *some* knowledge is required to ask good questions and pursue the answers we receive. But if we are honest about this we will admit that the kind of exhaustive expertise we typically expect in students is overkill. After all, children are wonderful and persistent questioners: recall the wisdom of H.C. Andersen's "The Emperor's New Clothes." Indeed, academics are invariably prone to making the mistake Gilbert Ryle called the Cartesian fallacy: assuming that "knowing that" must *always* precede and serve as a condition for "knowing how." No person who creates knowledge or uses knowledge to put bread on the table would ever be guilty of this fallacy. All apprentices or would-be performers learn on the job, yet as teachers we tend to over-teach or "front load" knowledge. So a good pedagogical rule of thumb would be: teach the minimum necessary to get the students asking questions that will lead to your more subtle goals.

We would do well, then, to think of our task as introducing the student to cycles of question-answer-question and not just question-answer—with one aim of a course being to make the student rather than the professor the ultimate initiator of the cycle. To postpone developing students' ability to ask important questions in the name of "mastery" is to jeopardize their intellect. Good judgment and aggressive thinking will atrophy if they must be endlessly postponed while professors profess. In any event, the most important "performance" in the liberal arts is to initiate and sustain good question asking.

A very mundane point about testing can be made out of this esoteric argument. We rarely assess students on their ability to ask good questions. Indeed, we rarely teach them a repertoire of question-asking strategies for investigating essential ideas and issues. It should become obvious to students through the demands of the course and our assessment strategies that question asking is central. Too often, however, our assessments send the message that mastery of the "given" is the exclusive aim, and that question asking is not a masterable skill but a spontaneous urge.

Principle #8. This principle follows from #7. The aim of the liberal arts is to explore limits—the boundaries of ideas, theories, and systems. To paint the starkest picture of the difference between a "liberal" and a "non-liberal" view of the disciplines, therefore, we might see our task as teaching and assessing the ability to gauge the strengths and weaknesses of every major notion we teach—be it a theorem in math, a hypothesis in science, or literary theory in English. We need to know whether students can see the strengths and weaknesses of "paradigms." This would include the limits of a theory not only within a subject but across dis-

ciplines, as when we apply the rules of physical science to the human sciences.

There is no novelty in this idea. I am invoking a notion about the liberal arts developed 30 and more years ago by Joseph Schwab at Chicago. He termed such a view of education the art of "eclectic," and I encourage you to return to his essays for numerous suggestions on how to help students explore the merits of sanctioned truths.

I fear that we no longer know how to teach science as a liberal art in this sense. When we make science merely abstruse and technical, we make it increasingly unlikely that non-scientists will profit from studying science enough to support intelligent science policy as adults. And we encourage science students to become too technical and insufficiently critical. I really think that the first years of study in college (and certainly throughout secondary school) have less to do with "mastering" science and more to do with orthodox algorithms—learning metaphysics instead of physics, sanctioned truths vs. the unstable results yielded by methods and questions that *transcend* the current results.

I know this weakness in our science students firsthand from my high school teaching days. My students did not understand, for example, that error is inherent in science and not merely the fault of immature students or poor equipment. (Many believe that when the "big boys and girls" do their measuring, the results are exact). Nor did many of them realize that words like *gravity* or *atom* do not correspond to visible "things" to be seen directly.

The point can be made another way. We still do a poor job of teaching and assessing the student's grasp of the history of important ideas. I know of no other method by which inappropriately sacred truths can be more effectively demystified. What questions was Newton, then Einstein, trying to answer? What did the first drafts of a history text look like, and why were they revised? To ask questions like these is to open up a new and exciting world for students. To be smug about our knowledge and to distance ourselves from "crude" and outdated theory is to ensure that we repeat the mistakes of our smug and parochial elders.

Consider the history of geometry, the very idea of which strikes many people as an oxymoron. Many college students are utterly unaware of the problems that forced Euclid to develop an awkward parallel postulate (which was instantly decried by his colleagues). So much for self-evident truths, that glib line found in superficial views of Greek mathematics! Further, most students are unaware that non-Euclidean geometries can be proven to be as logically sound as Euclid's; they have no idea how that result transformed epistemology for all fields of study.

The consequence of our failure to reveal to students the history of important ideas is twofold. For one, students easily end up assuming that axioms, laws, postulates, theories, and systems are immutable, even though common sense and history say otherwise. The second result follows from the first and does lasting harm to intellectual courage in all but our feistiest students. Students never grasp that "knowledge" is the product of "thinking"—thinking that was as lively, unfinished, and sometimes as inchoate as their own. One major reason for the intellec-

tual poverty in this country is that most students either become convinced they are incapable of being intellectual, or they are uninterested in it if it involves only arcane expertise in a narrowly framed subject.

Some practical implications for assessment? First, we should require students to keep notebooks of reflections on coursework, their increasing knowledge, and important changes of mind about that knowledge. Second, we should assess this work as part of the grade. I did so for many years and found it to be the most important and revealing aspect of the students' work. I also learned a lot about how their thinking evolved in a way that improved the courses I taught. Third, even the most technical training should ask students to do critical research into the origins of the ideas being learned so that students can gain greater perspective on their work. If we fail to do this, whether out of habit or rationalization that there is no time for such reflective work, we risk producing a batch of thoughtless students.

Principle #9. Number 9 extends the moral implications of #8. When we encourage narrow, unchecked expertise, we may unwittingly induce students to be dishonest about their ignorance.

I am not even talking about the more heinous crime of cheating, something we know is all too common. Rather, I am talking about the moral obligation of the liberal artist to emulate Socrates's trademark: his cheerful admission of ignorance. Alas, our students rarely admit their ignorance. One of our primary tasks should be to elicit the admission and not penalize it. But the student's willingness to take such a risk depends upon our doing so. It is then, as the "Meno" reminds us, that mutual inquiry and dialogue become possible because we are placed on equal moral footing as thinkers. More pointedly, our inclination to "profess" is always in danger of closing off the doors through which our students can enter the liberal conversation without excessive self-deprecation. So many of our students preface a wonderful idea by saying, "I know this sounds stupid, but . . . ."

Let our assessments, therefore, routinely encourage students to distinguish between what they do and do not know with conviction. Let us design scoring systems for papers that heavily penalize mere slickness and feigned control over a complex subject, and greatly reward honest admissions of ignorance or confusion. Or, let us go the next step and ask students to write a second paper in which they criticize the first one.

Principle #10. The 10th and last principle extends the point. Intellectual honesty is just one aspect of self-knowledge and the absence of self-deception. One of my favorite notions was put forward by Leo Slizard in talking about how to assess doctoral candidates. He argued that students should be assessed on how precisely and well they knew their strengths and limitations—and that it was a mistake to err greatly in *either* direction.

I am not arguing for professors to become counselors. I am arguing for them to improve students' ability to self-assess, and to make sure that accurate self-assessment is more than a pleasant exercise. It is an essential tool for ensuring that students have neither excessive

nor insufficient pride in their work, either of which closes off further intellectual challenges and rewards.

The inherent danger of scholarship is not so much error as blind spots in our knowledge—hidden by the increasingly narrowed focus of our work and the isolation that can then breed worse: arrogance. Excessive pride leads us not only to ignore or paper over our doubts but more subtly to be deceived about the uniqueness and worth of our ideas—we forget that it was a conversation in the coffee shop or reading an article that sparked the idea. A few collaborative assessment tasks, with some reflection on the role of each contributor, would provide useful perspective for everyone.

It follows that we should assess class discussions more than we do. We again fail to assess what we value when we make it possible for students to learn everything from only listening to us and doing the reading. I and many others have developed some good material for assessment (and self-assessment) of class discussions, and I encourage you to develop some methods of your own.

Which of course brings us back to Socrates. What the casual reader of Plato always fails to grasp—including some overly analytic philosophers, I might add—is that the dialogues invariably are about character, not "theories" of virtue, knowledge, or piety. The twists and turns of dialogue, the sparring with Sophists or young know-it-alls—ultimately all this is meant to show that character flaws, not cognitive defects, impede the quest for knowledge. It is one's *attitude* toward knowledge that ultimately determines whether one will be a liberal artist or merely proficient.

As Socrates repeatedly reminds us, we must love wisdom so much that we question our knowledge, even our pet ideas if need be. By extension, the more we gain confidence in our ideas, the more we must become vigilant about finding knowledge in unexpected places—and let others who seem incapable of it teach us something, as our students often do.

It is not a canon—of ideas or books—that defines the liberal arts, but a set of very hard-won virtues. Like all sophisticated dispositions, these liberal habits are typically only revealed when they are challenged. It is only when peer pressure is greatest, be it in the classroom with students or at conferences with our peers, that we learn who has the power to keep questions alive. The liberal arts, properly speaking, do not make you free; they *keep* you free. Wisdom—Socrates knew—reveals itself when persistent inquiry is threatened: externally by custom and "oh, *everyone* knows . . ."; and internally by the tendency to rationalize our own habits, beliefs, and fears.

How much do students really love to learn, to persist, to passionately attack a problem or task? How willing are they, like the great Indian potters of New Mexico, to watch many of their half-baked ideas explode, and start anew? How willing are they to go beyond being merely dutiful, perfunctory, or long-winded? Let us assess such things, just as the good coach does when he or she benches the talented player who "dogs" it.

We are then quite properly assessing not skill but intellectual character. It is to our detriment and the detriment of the liberal arts if we feel squeamish about saying and doing so. Let us "test" students

in the same way that the mountain "tests" the climber—through challenges designed to evoke whether the proper virtues are present. And if not present, the quality of the resultant work should seem so inadequate to the *student* that little need be said in the way of "feedback."

Let our assessments be built upon the distinction between wisdom and knowledge, then. Too subjective? Unfair? Not to those who have the master's eyes, ears, and sense of smell—tact, in the old and unfortunately lost sense of that word. For these traits are as tangible as any fact, and more important to the student's welfare in the long run. It is not the student's errors that matter but the student's response to error; it is not "thoroughness" in a novice's work that reveals understanding but awareness of the dilemmas, compromises, and uncertainties under the arguments one is willing to stand on.

If our testing encourages smug or thoughtless mastery—and it does—we undermine the liberal arts. If our assessment systems induce timidity, cockiness, or crass calculations about grades and the relevance of today's assignment, we undermine the liberal arts. If our assessments value correctness more than insight and honesty, we undermine the liberal arts. If our assessments value ease of scoring more than revealing to students the errors or tasks that matter most, we undermine the liberal arts. Let us ensure, above all else, that our tests do just what Socrates's tests were meant to do: help us—and our students—to distinguish the genuine from the sham authority, the sophists from the wise. Then we will have assessments that are worthy of our aims. ■

ASSESSMENT AND THE WAY WE WORK

by
Pat Hutchings

ASSESSMENT AND THE WAY WE WORK*

by Pat Hutchings
Director, Projects on Teaching
American Association for Higher Education

Twelve years ago, a still-damp Ph.D. in hand, I went looking for a job. The job market was tight—it was the worst of times for the humanities—and after living in the north woods of Wisconsin for a year to write my dissertation, I started sending letters to colleges in the upper Midwest—places I could go for interviews at my own expense. My companion in this was *Peterson's* guide: I started with institutions at the top of the alphabet and got all the way through *F* before I quit in despair.

But lo and behold, my despair was hasty. A few weeks later I found in the mailbox a letter from a place I only barely remembered from the *A*'s: Alverno College. They had a position, would I like to come down to Milwaukee and talk about it? . . .

I spent the next nine years at Alverno—and as an audience full of assessment practitioners would certainly know, those nine years included a lot of assessment.

I look back and see, more clearly than I did then, how much the presence of assessment at Alverno was linked to the ways we did our work there—*all* of us, faculty, staff, administrators, and students. And that's what I want to talk about today: assessment and the way we work.

I want to walk my way into that topic by recalling a few experiences at Alverno . . . particularly during my first semester, when the characteristics and conditions of the way we worked there were still new to me and therefore visible . . . and in some ways startling.

*Presented at the Fifth AAHE Conference on Assessment, in Washington, D.C., June 30, 1990.

There was, for instance, the matter of the syllabus. EnII4: Introduction to Literature was one of my first-semester assignments, and I set enthusiastically about the task of constructing a syllabus. Things were going along nicely when I began to realize that what I was doing—listing books and when they were to be read, as every syllabus I'd ever seen did—was not quite what my experienced colleagues were doing. I got talking with one of them in the hall about which of the way-too-many stories in the anthology I was using would work best, and I was asking her whether John Barth's "Lost in the Funhouse" wasn't a neat choice. Students would like that wouldn't they? And she said, Well, *she* certainly liked the story . . . and we should talk about Barth sometime. And as to the syllabus, well, let's see . . . what was I trying to *accomplish* during the class session in question?

Now I don't want to make myself sound greener than I was. But the fact is that my colleague's comment that day was one of those seemingly small things that one comes to see as a bench mark. Here I'd been thinking that the challenge of the literature classes I'd be teaching was to get the class *through* fifteen weeks of short stories . . . now I find that they're supposed to *learn* something. And following from that, that part of my job was to tell students what, as a result of all those stories read and discussions had, they would know by the end of my course . . . and beyond that, to give an account of how I and they would know whether they were indeed learning those things.

I need hardly say that these had not been the principles at work in my recently completed graduate courses. But they *were* principles that made sense. And, as I came to see, their good sense worked not only for students but for me. Getting clear in the way that assessment demands about what you want students to learn can be a long and frustrating experience. But, as another Alverno colleague said to me years later, it's also a great relief, down the line, to look at the foot-high stack of student papers on the corner of your desk and not to wonder, Why did I assign this one?

So that was the syllabus.

And then there was the matter of off-campus experiential learning. That's OCEL, another of my first-semester teaching assignments, this one with a team of faculty from various disciplines whose students were engaged in off-campus work where they were to apply and test out and refine what they learned in classes in their major. Once again I found myself in the midst of questions I'd never considered. Questions not just about the kind of learning that might be expected from an introductory literature class or even from the English major, but questions about how that learning would contribute to the student's ability to perform *beyond* the field of literature, even beyond the *campus*. What difference, for instance, would the fact that an English major had read (let us say) "Lost in the Funhouse," or any number of short stories, make in the way she did her work at a project at the Milwaukee Historical Society? What did all those discussions of and that writing about literature add up to in terms of her ability to make sense of the archival material she was working with? How did it affect her ability to work with professional colleagues?

I don't pretend even now (especially now) to have neat answers to those ques-

tions, but asking them was, again, an eye-opener. Moreover, because there were students in the OCEL seminar from a variety of disciplines, I found myself confronted as well with questions about "larger outcomes" that cut across disciplines and departments (the ability to learn from one's own experience was a key item) and about my responsibility for seeing that students achieved those outcomes . . . and *teaching* toward them.

And then we come to my writing lab, where, through various specified steps, students worked on papers for other courses. This lab was unusual, however, in that my job was not only to help the student do her best job on those papers and learn from the process of writing them, but also to collaborate with the instructor who had assigned the paper (a sociologist, let us say); to sit down with that sociologist at various points in the semester and come to a shared judgment about whether our mutual student was meeting our agreed upon expectations—*our* expectations, which were also the agreed upon expectations of the *college*, for which all of us as faculty were responsible . . . together . . . this, if you will, the corollary to the point above about cross-cutting outcomes: shared responsibility for those outcomes. Shared too—and not just incidentally—with the student, whose assessment of her own strengths and weaknesses as a writer was part of what I and her instructor took into account in evaluating her performance.

So, it was an interesting first semester. The second semester was pretty interesting too. In March, I was invited to sit down with my division chair to talk about how things were going, and to negotiate a contract for the coming year. What—I was asked—were my goals as a teacher? What progress had I made toward those so far? What evidence did I have of that progress? What had I learned from colleagues who visited my classes as peer evaluators? What would I be working on next year? And we looked at and talked about the criteria for successful teaching performance at the assistant-professor level as set forth in the faculty handbook, where I found, too, statements about the college's expectation that in the future I would be evaluated not only on the basis of effectiveness in my *own* classroom but in terms of my contributions to our collective effectiveness as teachers across the institution. Interestingly, I don't think I was particularly startled by that point back in 1978—it was completely consistent with everything else that was happening—but it seems more than a little surprising now. What an idea: that teachers should be responsible for one another's teaching!

In talking about performance review (as I learned much later to call it), the writing lab, the experiential learning seminar, that first syllabus, I'm zooming in very selectively on what is, of course, a much bigger, more complicated landscape. I mention these features of work at Alverno because they seem to me to get at the very heart of assessment—which is not, finally, a matter of methods or mandates but a way of thinking about learning:

- thinking of learning as the aim and even the test of teaching;
- recognizing and teaching for types of learning that transcend the individual course;
- assuming shared responsibility for that learning; and

- at the institutional level, thinking of learning as the end—the valued and rewarded end—toward which other things are consistently aimed.

Let me turn now from Alverno to a more general discussion of how these principles of assessment—for that's what I understand them to be—play themselves out in the way we work—alone in our offices struggling with a syllabus, in classes with students, with colleagues in the department, and across the institution. I want to talk about how assessment is connected to, helps us see, and maybe even changes, our work with students and one another and the conditions that make that work *work*.

Work With Students

The best account I've heard of how assessment changes the way we work with students comes from a faculty member from a community college, a teacher of writing, who wrote to the AAHE Assessment Forum to say, "Assessment means to me, asking whether my students are learning what I think I'm teaching."

As I say, I like that question very much; it puts in a nutshell many of the insights I was coming to at Alverno that first year. But there are other questions that might be cited here as well. They'll sound familiar to you, I think:

- What do I have to do to get an *A*?
- Why do we have to do this paper?
- Why don't you just tell us what you're looking for?
- Where is this course going, anyway?
- Do we have to know this?

Those are, of course, students' questions—and not, I'd like to propose, the questions of grumblers only, but of *good* students who quite reasonably want to know what's expected of them and where things are going; and they're questions, too, of students who could be a whole lot *better* if they understood the answers. They're questions that assessment speaks to very directly, as the following two examples show.

The first is a project supported by the Fund for the Improvement of Postsecondary Education (FIPSE) being undertaken by 10 (and over the next two years 20, and then 30) community colleges in California, all part of LARC, the Learning Assessment Retention Consortium that began several years ago to collaborate on the assessment of entry-level skills. This past year a number of LARC institutions got interested in looking not just at entering abilities but at outcomes down the line. And they wanted to look at "outcomes" in a way that took advantage of faculty's interest in and commitment to teaching. What they turned to is K. Patricia Cross's concept of Classroom Assessment.

I visited the group earlier this spring and got to hear them talk with one another about what happens when faculty and students engage in the small-scale, in-class assessment methods that Pat and her colleague Tom Angelo have proposed.

There was, for instance, the chemistry professor from Cuesta College who confessed to the group that though he was an experienced teacher (and obviously one who cared a lot about students), he was feeling like a beginner with this assessment stuff, not sure that he was "doing it right." . . . But the more he and his colleagues talked, the clearer it was

that something important was happening that was changing the way these faculty were working with students.

Like several others in the group, this professor was using the "one-minute paper," where students write briefly at the end of the class about what they understood that day, and what they're still not clear about. And he talked to the group about what he was learning about his students—about how they think about the problems he assigns, where they get stuck, where breakthroughs occur. "I just never knew this much about my students' thinking," he told us. Wanting to understand more, I asked him, "So, the power of Classroom Assessment is in giving you more information about students. Is that right?" And he paused and said, "Well, yes, it's information. But really, you know, the bigger difference is that I never thought to *ask* these questions before."

At the risk of putting words in this fine man's mouth, I can't help but note that he was having the same insight I had at Alverno as I worked at my syllabus that first semester—that as soon as you ask, as assessment does, Are students getting it?, you're also asking yourself to be much clearer about what that *it* is, what you expect, where you're aiming . . . and the whole enterprise gets more purposeful and focused. That line of thinking isn't an easy one; it's not one most of us were ever trained to undertake. But it sure can change the way faculty and students work together.

Faculty at SUNY-Fredonia (my second example) have come, I think, to a similar conclusion.

For several years now, with FIPSE funding, Fredonia has been assessing a new general-education program, and they've chosen to go the local route in all cases, with faculty-designed exams to cover a variety of cross-cutting, general-education outcomes. I think that's a right idea, and when I visited there several years ago it looked like things were going well, and so I was dying to read, as I finally did this winter, their final report to FIPSE. In that report they set forth what they've learned—and equally interesting, what they didn't learn.

They didn't, for instance, ever learn the Truth about their students. In fact, much of their work at the outset entailed discussion—and I dare say heated debate—about the soundness of this and that instrument, pilot testing, whether the results could be compared this way or that, what was valid, what not . . . with the result, as I say, that no Truth was learned. What faculty *did* learn was that whether the score was 37 or 43 (a debate that can go forever), students weren't doing as well in "reflexive thinking" as faculty thought they should. That is, students were taking and passing individual courses alright, but they weren't seeing connections; they couldn't put the pieces together.

The solution? No doubt there were (and should be) several. But interestingly, one of their next steps at Fredonia is more assessment. Faculty are now working to develop a portfolio approach to assessment that will give them more in-depth information about each student's ability to put the pieces together, but also—and here's the beauty of the thing—help the student *develop* that ability.

Portfolios are kept by students themselves and include work done through the four—or whatever—years. The first

piece, done at entry and used for advising, is "My History as a Learner," an essay where the student is asked (i.e., told that at Fredonia they will be *expected*) to self-reflect, to be conscious of what they do as learners. Intermediate pieces in the portfolio are, I believe, pulled from various course settings; the final entry is an essay in which students look back on all that work and are asked to put the pieces together, addressing, in particular, questions about how general education is connected to the major.

It is, as I say, a way of gathering very rich information about students. More important, it's a method that's for teaching first, assessment second . . . or that blurs the difference. Finally, it's a method that recognizes that the coherence of the curriculum has to reside not only in *our* neat schemes but in the student's ability to make connections and create coherence, an ability that can be attended to, focused on as an expectation, taught for, assessed, and improved.

These accounts of work being done in the LARC group and at Fredonia risk making things seem smoother than they are: People from the institutions in question may, I'm aware, be squirming in the audience. My intent is not to suggest that they've somehow got things righter than the rest of the world but to say instead that I *do* see in both projects (and in many more featured at this conference) some ways of thinking that have the potential to change the way we work with students . . . by making learning the test of teaching; by nudging us into a clearer view of that learning; by finding ways to communicate that view to students, to let them know, through statements in syllabi and through the assignments and assessments we give, what we expect, that we expect a lot, that we and they are going to meet those expectations.

By way of summary to this section on the way we work with students, I offer a comment that my colleague Barbara Wright, director of the AAHE Assessment Forum, made recently in response to a draft of something I was writing. I was getting heated up over the fact that the burden for learning so often falls so exclusively on students. "Students responsible for learning, . . ." Barbara wrote in the margin (I could imagine her eyebrow cocking). "Not a bad idea." It isn't, and assessment, if you will, agrees. It speaks to a shared responsibility between faculty and students. It calls on us to create conditions in which students can carry out that responsibility, where they don't have to ask, Where is this course going? What are you looking for? What am I learning here, anyway?

Work With One Another

Stories about how we work with students are, of course, also stories about how we work with one another. In fact, Pat Cross has remarked that something she didn't anticipate about Classroom Assessment—one attraction of which would be that it could be done alone—is that faculty no sooner get into asking the kinds of questions the LARC faculty were excited about, than they want to talk to and hear from colleagues who, it turns out, are also interested in those questions . . . and dying to talk about them.

What we see in those cases and on scores of campuses is one of assessment's most powerful and in many cases immediate effects: the prompting of a conver-

sation about student learning that otherwise takes place not at all, or only on the most sporadic basis, or in ways that have little chance of finding their way into institutional decision making.

I'm thinking here of a hundred stories I've heard. From a faculty member in the art history department at the University of Tennessee-Knoxville, for instance. She said it was only when faced with assessment—in this case a state-mandated requirement to examine the outcomes of the major—that faculty in her department "for the first time sat down and talked with one another," not altogether happily, mind you, about what goes on in their respective classes, what's covered and (one hopes) uncovered, what the aims of various courses are, what assignments are given . . . all of that otherwise behind-the-doors stuff that assessment opens onto.

And I think about a man I met during a visit to the University of Connecticut. An English professor, far from convinced that assessment was the best thing in town, who said he's stuck with the work of the assessment committee because it is "the one place on campus where there was a serious conversation about student learning going on."

And I think about Harvard. Some of you may know the Harvard assessment story from the NY *Times*—the front page no less—where a report on the Harvard Assessment Seminars appeared a few months ago. Or better yet you may know it from the May *AAHE Bulletin*, where editor Ted Marchese interviews the convener of those seminars, Richard Light.

It's work that goes back to 1986, when Harvard president Derek Bok urged in his book *Higher Learning* that every college "study the learning process and assess the effects of its programs." To that end, on his own campus, he asked Richard Light, a faculty member in the Kennedy School and the Graduate School of Education, to convene a seminar on undergraduate learning—to which more than 100 faculty and administrators were soon drawn.

Interestingly, the story of the seminars, which were held over three years and still go on, is very similar to that of the LARC Classroom Assessment project. Though there have been "findings" in the Harvard picture, its seminars are best described as "faculty inquiries into student learning."

For example, participants took on questions about the impact of the size of groups in which students study, which led to several experiments conducted by faculty with their own students that showed that students who study in groups of four to six do better academically than students studying alone.

They looked, too—at the urging of student participants in the seminars—at questions of gender, and discovered that women students at Harvard have an experience rather different from men's.

In the *Bulletin* interview, Ted Marchese says to Dick Light, Well, now, that's all very nice—these findings about gender and about study in groups—but it certainly isn't news. Light has a great comeback (exactly the one Ted was after, I suspect): "Newness," Light replies, "is hardly the goal here; we're after locally useful information and small but steady increments of improvement." He knows, he says, "that similar findings, some from earlier decades, exist in the library, but there is a power, an immediacy, that

comes out of your own discoveries." The upshot is that long-time, long-tenured faculty are now talking to one another about how students learn, and doing things as a result that they didn't do two years ago—among them Pat Cross's "one-minute paper."

The point? It's not that Harvard has made some quantum leap forward in quality; it's that assessment has helped create an occasion to take up, collectively, questions about learning and the conditions under which it can occur best.

At Harvard, UTK, UConn, and scores of other campuses, assessment has changed the way we work by getting us talking to one another, across all kinds of lines and boundaries, coming to clearer, more collective visions of our aims and purposes, asking questions together about whether we're achieving those purposes and how we might do better.

Alverno is an interesting footnote here: By my lights, one of the most important things Alverno has done—more important than any single assessment innovation the college is admired for—is to set aside Friday afternoons for faculty to work together on questions of teaching and learning . . . questions prompted and then illuminated by assessment.

The Culture of Assessment

Thinking about assessment's impact on the way we work with students and with one another, we come to a deeper question, one about the kind of culture—the conditions—in which our best work can go forward. What kind of culture is that? What habits of mind and deeper structures make the work called for by assessment work?

In some ways I think this Assessment Conference is an answer to that question. As I've gone from session to session over the last three days, I've been struck that underlying what many of you are doing in the name of assessment is . . . what shall I call it? . . . a view of the world, a subtext, a vision of what our institutions might be like, a set of educational values. Assessment is not, I've come to believe, just a set of methods; not a technology to be plugged in anywhere toward any end.

It is (and this was my point in beginning with Alverno) an enactment of a set of beliefs about the kind of work that matters on our campuses. Here are three:

1. That teaching is to be taken seriously and rewarded.

The issue of rewards may appear to be the more pointed one here, but let me pause for a moment on "taking teaching seriously"—the title, not incidentally, of a keynote address Pat Cross delivered at AAHE's 1986 National Conference. It is, I think, no accident that the spokesperson for "taking teaching seriously" also speaks for Classroom Assessment. Assessment is a way of enacting a greater seriousness about teaching; it also challenges us, I think, to look harder at what it would mean to be really serious about teaching.

I think, for instance, of a piece that appeared a year and a half ago in *Change* magazine, entitled "Claiming Ourselves as Teachers." In it, Diane Gillespie, a faculty member from the University of Nebraska, argues that public discourse about teaching is scarce and indeed thin because teaching itself has been privatized, pushed off into the margins of

work. To be caught talking in public about teaching is, she says, "like discovering at a formal dinner that you're eating someone else's salad."

Gillespie's account of things is no doubt more true on some campuses than others—more likely to be true, one supposes, on research university campuses like her own. But what's worrisome is the possibility that it might be true even on the many more campuses that identify themselves as "teaching institutions," a label that seems too often to mean an absence of research but not necessarily the *presence* of sophisticated conversation about or inquiry into teaching and learning.

What would that presence look like? It would look, I think, a lot like assessment at Harvard and SUNY-Fredonia and the LARC institutions: more conversation about teaching and learning, more *well-informed* and *collective* conversation that has a home in the institution.

The amount and character of discourse about teaching is one sign of the seriousness with which it is regarded. Another is rewards. A colleague here in D.C., Christine Young, recently wrote to me, "Until evidence of teaching effectiveness is taken seriously as a criterion for hiring, promotion, tenure, and merit raises, those faculty members who take teaching (and therefore assessment) seriously may continue to function at the margins."

That's a hard statement and a true one. If you follow assessment to its logical conclusion and ask what it's after—more student learning is a simple version of the answer—you come precisely to Young's point. Assessment will be an add-on, a marginal thing, until we reward teaching.

Interestingly, a method that's caught on in a big way (it was on this conference program prominently) for purposes of student assessment may be promising as well when it comes to evaluating teaching. I'm talking about portfolios, in this case "the teaching portfolio." We don't yet have the perfect formula for exactly what materials, what documents, could most usefully be displayed in a teaching portfolio. But the *fact* of the thing, a collection of evidence of teaching effectiveness—work samples—put on the table in the context of faculty evaluation, is already powerful. And the act of assembling evidence—like the fact of assessment itself—is likely to prompt better thought about how good teaching can be known and documented and rewarded, and what the institution *means* by "good teaching."

The issue of rewards for teaching gets a further, interesting spin when looked at in the light of assessment. As things now stand, all the rewards run to the individual. Assessment, with its principle of collective responsibility for student learning, implies a need for collective rewards. And, in fact, 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. That may be the proverbial exception that proves the rule; it's also very suggestive in thinking about assessment and how it might shape—reshape rather radically in this case—the conditions in which we do our work.

2. That assessment not only values teaching, it has a view of learning.

If you look at what's being done in the name of assessment on many campuses

today, you see an emerging view of learning: What matters in this view is not just what students know but what they can *do* with what they know. What's at stake is the capacity to *perform,* to put what one knows into practice. And this focus on performance implies asking not just how much (seat time or credit hours) but how *good* (where all parties understand what constitutes "good").

This conception of learning as performance has yet a larger aim. Assessment presumes a kind of learning in which students—knowing what is expected of them—can with practice over time become their own best assessors. The object here is graduates who know their own strengths and weaknesses, can set and pursue goals, who monitor their own progress and learn from experience. There's considerable evidence now that students who are self-conscious about their processes as learners are better learners, that they learn more easily and deeply, and that their learning lasts. The fashionable label for the skills in question here is "metacognitive," but whatever you call them they represent a kind of learning that speaks to a belief that learning is personally liberating, self-empowering, and for *all* students.

3. Finally, that the culture of assessment is one in which we not only aim toward a particular kind of learning but hold ourselves accountable *for it; where accountability is not a dirty word (what "they" want), but part and parcel of the way we work.*

Most of the talk about accountability so far has been about that which runs from the institution to the state. A few years ago you heard this talked about mainly in terms of reporting requirements, comparing of scores, and so forth. More recently, and largely because of the work of Peter Ewell, institutions are finding more constructive ways to think about this kind of accountability. But those that are taking assessment seriously have a different and additional slant on accountability, as well. While recognizing obligations to external publics, they want to look inside to deeper-running responsibilities.

A first of these runs from the institution to its students. Accountability here means delivering an education equal to that promised in recruitment, to the student's investment (not only of money but of time and effort), and to the demands of the student's postcollegiate life.

A second kind of accountability implied by assessment is that of students for their own learning. A number of institutions are now teaching students to "self-assess," to diagnose their own progress, to take responsibility for it, to ask the "what-it-adds-up-to" question of *themselves* as learners.

Third, and most important, there's the accountability educators have to one another, in our teaching and related work, on mutually agreed upon purposes and promises. This is the professional obligation that goes with the autonomy that faculty have traditionally enjoyed; it invites many ways of working with students, but asks also that the work each of us does individually contributes to the larger aims we've agreed upon together.

Seen this way, improvement and accountability—the oft-cited tension behind assessment—are, if ends of a continuum, ends that come around and meet.

Themes for the Future

By way of summary, let me begin—presumptuously perhaps—by making a point on behalf of all of us here who are faculty members . . . one that comes, in fact, from a letter I recently received from a faculty member from SUNY-Fredonia who's been involved in the work I described earlier. "I am concerned," Patrick Courts writes, "that assessment has become a self-regressive project concerned primarily with itself—concerned with more and better assessment but little beyond that. Personally, I have little use for it unless we allow the voices of those assessed to be heard." That's a right and timely caution. It's one AAHE tried to keep in mind when putting this conference together; it's one to keep a constant eye on in the future.

Second, we need more leadership for assessment. From top-level administrators we need the kind of leadership that expresses itself through rewards and incentives. And those need to be for assessment itself (designing and doing it) but also, more important, for the ends we want assessment to serve: better teaching and learning. But we need leadership that goes beyond "good management" here. Too often assessment is seen as an administrative problem to be solved with administrative responses. What we need is educational leadership. Ideas about learning. And we need that from administrators and from faculty and, yes, from students too.

Third, we need assessment that follows from the ways we think about learning, the kinds of learning we value. We need to get clear about what those are. And if, as I've argued, we value not just disembodied facts and knowledge but what students can do with what they know, we need assessments that call for and document those abilities. This, I take it, is what Grant Wiggins has in mind when he talks about "authentic assessment," assessment "composed of tasks we value," assessment that is "standards-setting."

Finally, we need more than assessment to get where we need to go—that is, to more and better student learning. This, in my view, is the single most important point to be learned from Alverno. Over and over I've been asked, Has assessment actually improved things at Alverno? How do you know? That's a complicated question, but *one* relevant answer is that the growth in student learning that Alverno has been able to document is a function not of any *one* thing but of a powerful culture of learning: a consistency and clarity of purpose; teaching aimed at that purpose; a sense of responsibility to students (and *by* students); a sophisticated, institutionwide conversation about learning; and a view of teaching as a valued professional activity.

Assessment—I think my Alverno colleagues would agree with me here—is essential to all of the above, but it's also part of a bigger picture. What's at stake is not assessment but the larger ways we work. ■

About AAHE

The American Association for Higher Education (AAHE) is a national organization of individuals dedicated to improving the quality of American higher education. Its members include faculty and administrators as well as educational policy makers. Through AAHE conferences, publications, and special projects—like the AAHE Assessment Forum—members can acquire both the "big picture" and the practical tools to increase their own effectiveness.

The benefits of membership in AAHE include subscriptions to *Change* magazine and *AAHE Bulletin*; discounts on registration for the annual National Conference on Higher Education and other convenings; discounts on selected non-AAHE publications *(The Journal of Higher Education* and *ASHE-ERIC Higher Education Reports)*; and more.

AMERICAN ASSOCIATION FOR HIGHER EDUCATION

AAHE members receive free the *AAHE Bulletin* (ten issues/year) and *Change* magazine (six issues/year); significant discounts on conference registration; special rates on selected non-AAHE subscriptions; Hertz car rental discounts; and more. To join, complete this form and send it to AAHE, One Dupont Circle, Suite 600, Washington, DC 20036-1110.

Membership *(choose one)*

Regular: ☐ 1 yr, $70 ☐ 2 yrs, $135 ☐ 3 yrs, $200
Retired: ☐ 1 yr, $45 Student: ☐ 1 yr, $45
(For all categories, add $8/year for membership outside the U.S.)

Caucuses *(you must be an AAHE member; choose same number of years as new membership above)*

Asian/Pacific American:	☐ 1 yr, $15	☐ 2 yrs, $30	☐ 3 yrs, $45
Black:	☐ 1 yr, $15	☐ 2 yrs, $30	☐ 3 yrs, $45
Hispanic:	☐ 1 yr, $25	☐ 2 yrs, $50	☐ 3 yrs, $75

Name _____

Position _____

Institution/Organization _____

Address _____

City _____ St _____ Zip _____

☐ Bill me ☐ Check enclosed *(payment in U.S. funds only)*

About the AAHE Assessment Forum

A special project of AAHE and funded by FIPSE, the AAHE Assessment Forum was established in 1987 to help coordinate and advance thoughtful assessment. The Forum offers a variety of services and resources, including the annual AAHE Conference on Assessment; commissioned papers and conference presentations; national directories; an in-house assessment resource library; assessment resources via ERIC; assistance with consulting needs; and more.

For more information about AAHE or about its Assessment Forum, contact AAHE, One Dupont Circle, Suite 600, Washington, DC 20036-1110; ph. (202) 293-6440, fax (202) 293-0073.

Forum Resources

Assessment 1990: Understanding the Implications **$12.00**
From the 5th AAHE Conference on Assessment in Higher Education
- "Streams of Thought About Assessment" by K. Patricia Cross
- "The Truth May Make You Free, But the Test May Keep You Imprisoned: Toward Assessment Worthy of the Liberal Arts" by Grant Wiggins
- "Assessment and the Way We Work" by Pat Hutchings

Assessment 1990: Accreditation & Renewal **$10.00**
From the 5th AAHE Conference on Assessment in Higher Education
- "Assessment and Accreditation: A Shotgun Marriage?" by Ralph A. Wolff
- "Assessment as a Tool for Institutional Renewal and Reform" by Alexander W. Astin

Time Will Tell: Portfolio-Assisted Assessment of General Education (1990) **$9.00**
By Aubrey Forrest and a study group that included fifteen experts in the field. A comprehensive guide to the implementation and use of student portfolios to assess general-education outcomes at individual and program levels.

Assessment Programs and Projects: A Directory **$ 12.00**
Edited by Jacqueline Paskow; updated by Elizabeth A. Francis. (1987, 1990) Concise descriptions of thirty assessment projects implemented on campuses across the country—including purposes, key features, strategies and instruments, and impact.

Behind Outcomes: Contexts and Questions for Assessment (1989) **$ 5.00**
This paper sets forth nine areas of inquiry for assessment that gets "behind outcomes," with appropriate methods for addressing each area and resources for further work. By Pat Hutchings

Three Presentations—1989 **$10.00**
From the 4th National Conference on Assessment in Higher Education
- L. Lee Knefelkamp, "Assessment as Transformation"
- Peter T. Ewell, "Hearts and Minds: Some Reflections on the Ideologies of Assessment"
- Rexford Brown, "You Can't Get There From Here"

Three Presentations—1988 **$10.00**
From the 3rd National Conference on Assessment in Higher Education
- Alexander Astin, "Assessment and Human Values: Confessions of a Reformed Number Cruncher"
- Linda Darling-Hammond, "Assessment and Incentives: The Medium Is the Message"
- Robert H. McCabe, "The Assessment Movement: What Next? Who Cares?"

***Three Presentations—1987** **$ 5.00**
From the 2nd National Conference on Assessment in Higher Education
- Lee S. Shulman, "Assessing Content and Process: Challenges for the New Assessments"
- Virginia B. Smith, "In the Eye of the Beholder: Perspectives on Quality"
- Donald M. Stewart, "The Ethics of Assessment"

Resource Packet II: Six Papers $25.00

- "Acting Out State-Mandated Assessment: Evidence From Five States," C. Boyer and P. Ewell
- "Assessing Student Learning in Light of How Students Learn," J. Novak and D. Ridley
- "Faculty Voices on Assessment: Expanding the Conversation," P. Hutchings and E. Reuben
- "Feedback in the Classroom: Making Assessment Matter," K. Patricia Cross
- "Standardized Tests and the Purposes of Assessment," J. Heffernan
- "An Update on Assessment," (AAHE *Bulletin*, December, 1987), P. Hutchings and T. Marchese

***Resource Packet I: Five Papers $8.00**

- "Assessment, Accountability, and Improvement: Managing the Contradiction," P. Ewell
- "Assessment and Outcomes Measurement: A View From the States," C. Boyer, P. Ewell, J. Finney and J. Mingle
- "The External Examiner Approach to Assessment," B. Fong
- "Six Stories: Implementing Successful Assessment," P. Hutchings
- "Thinking About Assessment: Perspectives for Presidents and Chief Academic Officers," E. El-Khawas and J. Rossmann

*These publications are available only as photocopies.

ORDER FORM

Name _____

Position _____

Institution _____

Address _____

City _____ State _____ Zip _____

FORUM RESOURCES

No. of Copies	Title	Total
____	Assessment 1990: Understanding the Implications ($12.00)	_____
____	Assessment 1990: Accreditation & Renewal ($10.00)	_____
____	Time Will Tell ($9.00)	_____
____	Assessment Directory ($12.00)	_____
____	Behind Outcomes ($5.00)	_____
____	Three Presentations - 1989 ($10.00)	_____
____	Three Presentations - 1988 ($10.00)	_____
____	Three Presentations - 1987 ($5.00)	_____
____	Resource Packet II: Six Papers ($25.00)	_____
____	Resource Packet I: Five Papers ($8.00)	_____
	Delivery	_____
	Resources Subtotal	_____

Orders under $25 must be prepaid; orders of $25 and over must be accompanied by payment or purchase order. Make checks payable to **AAHE Assessment Forum**. Book rate postage is included in prices quoted. Allow four weeks for delivery. For first class, UPS, or foreign delivery, add $1 for 1 copy ordered; $1.50 for 2-4, and $2 for 5-20 copies. No refunds or exchanges.

JOIN AAHE

Members of the American Association for Higher Education receive the *AAHE Bulletin* (10 times per year); a free subscription to *Change* magazine (6 times per year); significant discounts on National Conference on Higher Education registration and special publications; and many other benefits.

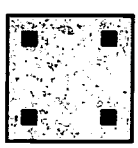
Regular membership:
- ☐ one year $70
- ☐ two years $135
- ☐ three years $200

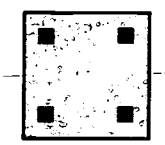
Special memberships *(one year):*
- ☐ full-time student $45
- ☐ retired person $45

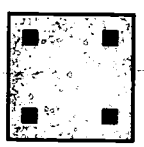
For all categories, add $8 per year for membership outside the U.S.

Membership Subtotal _____

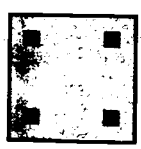
TOTAL ENCLOSED _____







AAHE ASSESSMENT FORUM
American Association for Higher Education
One Dupont Circle, Suite 600
Washington, DC 20036-1110
Ph (202) 293-6440 Fax (202) 293-0073



U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

Reproduction Basis

☑ This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☐ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").