

I am very pleased to be invited to be the main address of a company of such a distinguished group of professors.

USING ASSESSMENT TO IMPROVE INSTRUCTION

K. Patricia Cross

Harvard Graduate School of Education

On this quasi-spring day in Washington, it is tempting to compare the cycles of educational reform, that rise periodically to the top of the education agenda, to the seasons of the year. First comes the budding of ideas, quickly bursting forth into full-blown reform reports, vying with one another to produce the biggest, most colorful bloom in the garden of recommendations. Mild showers of criticism falling on the reports seldom dampen spirits. Instead, like spring rains, they refresh and clear the air for further discussion and debate.

Some seeds of reform fall on fertile soil and grow and are nourished. Others are buried under heaps of fertilizer, spread liberally and without differentiation across the landscape.

Yes, spring is the time for planting seeds, but it is also a time for weeds to grow and the sap to rise. Summer brings more light, greater maturity, and stronger roots. The crops that will be harvested in the fall depend as much on the care and skill of the gardener over the summer as on the quality of the seeds that were sowed so liberally in the spring.

Prepared as the DeGarmo Lecture for the Society of Professors of Education. Presented at AERA Annual Meeting, Washington, D.C., April 20, 1987.

We are now entering the summer growth period for the school reforms that started with the publication of A Nation At Risk in 1983. Between 1983 and 1985, nearly 300 task forces were appointed throughout the fifty states, to make recommendations about what should be done to improve the quality of education in America. This first generation of reform reports was heavy on top-down reform. Governors and state legislators talked a lot about what they could do to bring about reform in the schools. The notion of "assessment" was among the hardiest of the fruits of their labors, and for almost every level of education, from grade school through college, the issue ^{is} ~~was~~ how we--the American public--^{can} ~~could~~ find out what students ^{are} ~~were~~ learning or not learning as they ^{make} ~~made~~ their way through the American school system.

We are now into the second generation of reform reports. In the last two years--1985 to 1987--forty new reports on education have been directed to national audiences by organizations of every conceivable stripe, including those representing teachers and teachers' unions, principals and chief state school officers, foundations, state and national departments of education, colleges of teacher education, and deans of schools of education. ^{These} Recent reports differ from the first generation reports in one very significant respect. They are less concerned about what legislators can do to improve education and more concerned about what classroom teachers can do to improve the quality of teaching and learning. The second generation of reform reports advocates bottom-up reform.

In the best of all possible worlds, top-down reforms would meet bottom-up reforms gently and productively. It is with that hope in mind that I have entitled today's remarks, "Using Assessment to Improve Instruction." But I want to identify myself with the bottom-up reformers. I believe that in the final analysis, teachers who meet students in the classroom day after day hold the ultimate key to excellence in education. We should not expect to have excellence in learning until we have excellence in teaching, and we should not expect to have excellence in teaching until teachers feel responsible for the learning of their students.

Making teachers responsible for the learning of students is a heavy burden, one that we as professional educators have been reluctant to accept. We find it hard to commit ourselves to establishing the connection between teaching and learning. While we talk easily of teaching and learning, we are generally uncomfortable talking about teaching for learning, and we find it downright difficult to ask what teachers can do to cause learning.

In part, our reluctance to talk about causal links between teaching and learning is the result of our research methods, which have depended heavily on correlational studies. Much of the work on effective teaching, for example, searches for correlations between teacher characteristics and student learning. These studies shed only a dim light on what we might do to improve teaching, ~~In part this is~~ because many of the teacher characteristics that have been the subject of repeated correlational studies are quite resistant to change. There is

nothing educators can do, for example, to change age or gender or years of experience. Ben Bloom (1980) calls these "unalterable variables," and he urges researchers to gain more insight into teaching for learning through studying "alterable variables"--variables that can be changed by educators.

Researchers are now ~~beginning to~~^{ing} probe for causal links between what teachers do and what students learn. We are looking more closely at the process of teaching and learning. This means that attention has shifted from the study of teachers to the study of teaching. The challenge now seems to be to identify linkages between the process variables of teaching and the outcome variables of assessment. A report issued by the Education Commission of the States (1986, p.32) puts it well when they warned~~d~~ that "Assessment should not be an end in itself. Rather it should be an integral part of an institution's strategy to improve teaching and learning...."

Most people, I think, assume that assessment will improve instruction by documenting the strengths and weaknesses of student performance. Presumably, teachers will then use the results of assessment to take appropriate action. In higher education ~~especially~~, however, "taking appropriate action"^{usually} means making collective decisions about what is taught, i.e. about the curriculum. It rarely means doing anything about how it is taught. College teachers, especially, learn to profess disdain for the skills of teaching, and few have any training for their chosen profession. Fresh from graduate school and armed with voluminous and intricate knowledge of their specialty, they stand before a freshman class with little understanding of how students

learn. Students can and do learn without teachers, of course. Sometimes they learn in spite of teachers. But how students are taught is all-important. It makes the difference between a lifelong learner and a grade grubber, between enthusiasm for learning and indifference to it, between an educated society and a credentialed one.

If we are to improve the quality of teaching, we must find ways to restore pride to the profession of teaching and to instill respect for the skills and artistry of great teachers. I have thought a lot about this as I have pondered the suggestions for reform in higher education. My conclusion is that reform that calls for greater teacher responsibility, change as close as possible to where teaching and learning actually take place, and the empowerment of teachers to do the best job they are capable of doing, is the best route to the improvement of education.

To that end, I have proposed the development of a new set of skills that might be called "classroom assessment." I have also used the term "classroom research," ~~depending on the nature of the study~~. The purpose of classroom assessment is to help teachers evaluate their own effectiveness as teachers, through the systematic collection of information about student learning in their own classrooms. My hope is that through classroom assessment, teachers will become intellectually involved in studying the impact of their teaching on student learning.

Classroom assessment is not a new idea, but because it addresses so many of the problems that are surfacing in both secondary and postsecondary education, it may be an idea whose time has come. In the first place, the encouragement to

monitor learning and to study it through classroom assessment provides an intellectual challenge to teachers grown tired and bored with the repetition of doing the same old things in the same old ways. Teachers are often aware that some approaches are not working very well, but there seems to be little energy or reward for trying something new. The reward for classroom assessment is more intrinsic than extrinsic; it is driven by curiosity and involvement in seeking answers to the questions that teachers have. My hope is that, for some teachers, classroom assessment will offer an infusion of energy and the potential for experimentation and self-improvement.

A second reason for encouraging classroom assessment is to provide useful knowledge about effective teaching to the average classroom teacher. The criticism is often made that teachers fail to use the results of educational research to improve their teaching. For the most part, I think that is true. The problem is not really that practitioners do not or will not apply the findings of research to their practice, it is that they cannot. Most educational research is searching for generalizations across an infinite variety of teachers, students, and subject matters. In their search for general laws of learning, researchers hold constant or rule out the specific conditions in any particular classroom, *-- which, of course, is precisely what the CR teacher is trying to do in* ~~This makes the findings generalizable, but difficult to apply to any specific subject matter or classroom.~~ What a classroom teacher really wants to know is, What is happening in my classroom, given my students, and my subject matter? Classroom assessment is by definition situation-specific, and the intellectual involvement of the teacher in formulating questions

that matter to him or her is essential.

John Dewey (1929, p.19) wrote almost sixty years ago that "no conclusion of scientific research can be converted into an immediate rule of educational art." In the years since, we have only verified the notion that inspired teaching is more than the sum of research findings. In order to use research, teachers need to understand the findings in the context of their own teaching.

For college teachers, classroom research offers an absorbing alternative to discipline-based research. Most college teachers--70 percent according to a recent Carnegie survey (Chronicle of Higher Education, December 18, 1985, p.26) are more interested in teaching than in research. Yet, recent studies (Schuster and Bowen, 1986) are finding widespread pressures on ~~faculty members to publish, even on the part of institutions~~ *(even those in teaching institutions)* ~~historically committed to undergraduate teaching.~~ Since articles are being published today in some 40,000 scientific journals at the rate of one every 30 seconds, 24 hours a day, 365 days a year, it seems unlikely that college teachers who reluctantly enter this gate as the route to tenure are going to contribute much to our knowledge or to their own morale and productivity as teachers. Teachers using their classrooms as laboratories for their own investigations of how to cause greater learning may, however, contribute a great deal to their own understanding as well as to departmental discussions about the improvement of teaching.

The new evidence in educational research suggests that teaching effectiveness is best studied in the context of a

particular subject matter. What works in history probably won't work in math or studio art. As a matter of fact, teachers of piano may find research on teaching violin or French horn quite unrelated to their particular interest in teaching music. This doesn't mean that research on teaching and learning is so fragmented that no generalizations exist. It does suggest that we need all the researchers we can get to study effective teaching across hundreds of subject matters. Faculty who understand their discipline and who are in the best position to observe student learning, seem to have an excellent chance of making a contribution to knowledge as well as to the improvement of their own teaching.

Finally, classroom assessment compliments institutional and statewide assessment programs by starting from the bottom-up. If the ultimate aim of any recommendation for educational reform is to improve learning, there is considerable merit to starting the reform as close as possible to the scene of action. Although I think the assessments underway now will be helpful in providing some long-range perspective on student learning, institutional assessment or statewide assessment will probably not be very helpful to the average classroom teacher. Classroom assessment needs to accompany institutional assessment if improvement of instruction is the goal.

That moves us to the question, What is classroom assessment and how do you do it? For those in teacher education there is also the question, What is classroom assessment and how do you teach it to anyone planning to enter the profession of teaching? But that is another speech for another time.

For the present, my premise is that if teachers are to improve their impact on learning, they need to become systematic and insightful observers of student reactions to their teaching. This represents an important reversal of the usual approach to the improvement of teaching in which we are asking teachers to concentrate on their own behavior. It is not only quite difficult to observe oneself in action, but if students are not responding, it makes no difference how perfectly the teacher is teaching.) ^{CR helps to} to Morin p16.

Student reactions to teaching may be measured along a number of dimensions--cognitive, affective, attitudinal. Although we usually think of measuring ^{learning via} ~~reactions to teaching on~~ achievement tests, high school and college students are also quite capable of providing useful self-reports, especially if they are trained to observe their own learning in response to teaching.

Student evaluations of college teachers are common now in most colleges and universities, and required in many (Seldin, 1984). The problem with student evaluations isn't what most faculty members might think, i.e. that students are not good or reliable judges of teaching. From hundreds of studies of student evaluations of teaching, we can conclude that students generally know a good teacher when they see one. While there is frequently a range of opinion in any given classroom, when researchers go onto a campus and ask students, faculty, and administrators to identify the best teachers, it doesn't take long to come up with a list that shows high agreement.

Moreover, the evidence suggests that student ratings have good validity, i.e. that students tend to rate most highly those

courses in which they learn the most. Centra (1977) found correlations in the .60s and .70s between scores on final exams and student ratings of "overall teaching effectiveness" and "value of the course." While one can still find reported correlations ranging from negative to high positive, the tilt of comprehensive, well-designed studies--and the more recent meta-analyses (Cohen, 1982)--is clearly toward significant positive correlations between student achievement and positive course ratings.

Students also tend to be reliable and relatively unbiased raters. There is no evidence to support the myths that popular teachers are mere showmen, that the mature perspective of alumni will find virtues in the professors that were not apparent ten years earlier, that teachers who are tough graders will receive low ratings, or that agreement on the identity of good teachers is difficult to achieve (Gaff and Wilson, 1971).

No, the problem with student evaluations is not so much that they lack validity, as that they are not very helpful to teachers. The reasons, I think, are two: First, evaluation usually comes at the end of the semester when it is too late to do much about correcting problems--at least for that class. And second, most student rating forms are designed to evaluate rather than to provide the type of information that is most helpful to teachers. Students can be very helpful, especially if they are oriented to becoming careful observers of their own learning reactions to teaching.

Teachers can collect information about student reactions through the use of a wide variety of "feedback devices."

Feedback devices are measures that tell teachers what students are feeling, thinking, or learning. They may be simple, or complex, experimental or traditional, quantitative or qualitative, self-report or tests of achievement or measures of attitude change. Most teachers have developed some sensitivity to glazed eyes and puzzled frowns, but those are extremely crude feedback measures. Furthermore, teachers rarely check the validity of their interpretations of facial expressions and body language. Students, however, are sufficiently aware of these reactions to teaching that they have developed their own labels. MEGO is shorthand for "My eyes glaze over," and "Intra-ocular effects" means "results so clear they hit you right between the eyes." MEGO and intra-ocular effects are feedback measures that sensitive teachers can observe, but the point of classroom assessment is to make the collection and integration of feedback more systematic and useful.

One method of getting feedback on what students are learning is to ask them. But teachers need to learn to ask questions that provide good information about the quality of student learning. Researchers find that the questions asked by teachers in class are typically quite low in cognitive demands, usually calling for low-level factual knowledge, and usually not designed to provide the teacher with insights into student learning. Even when higher-order questions are asked, research shows that the chances are about 50-50 that teachers will receive and accept answers that do not match the cognitive level required (Berliner, 1984).

These findings from research on the process of teaching through questioning might simply be called "teaching tips," in

the sense that research shows that skillful questioning results in better student learning. But there is also a good opportunity for classroom research in these observations. A teacher might, for example, tape record a class session and later tally, for her own information, the type of questions asked and the type of responses accepted. Or she might involve students in the research, asking them to select any five-minute period in the class, during which the student would tally the cognitive level of questions and answers. Not only would such a scheme alert students to higher-level cognitive goals, but a comparison and discussion of three or four of the five-minute tallies might serve as a stimulus to class discussion about the nature of learning.

Let me give two other examples of simple feedback devices before moving on to some more complex designs. A physics professor at the University of California finds great value in what he calls "Minute papers." Four or five times a semester he stops class one minute early to ask students to write the answers to two questions: What is the most significant thing you learned today? What question is uppermost in your mind at the conclusion of this class session? He says this simple feedback device provides very useful information on how well students are understanding and whether there are important questions to which he should respond. Fred Mosteller, Professor of Statistics at Harvard, has devised a modification of this feedback device in which he asks students to write a couple of sentences at the end of the class period on the "muddiest point" in the lecture. This procedure, worked out with the involvement of his undergraduate

students in Stat-100, enables him to prepare a handout, or a clarification for the next class session, or to rework his explanation for the next time the class is taught.

Other feedback devices are inherent in examinations, especially midterms, but valuable information on student learning is frequently ignored in the single-minded effort to assign a grade. With some imagination, questions can be designed to reveal, not only which questions are missed, but why. Some teachers get even more feedback by asking students to evaluate the test questions. ~~Was the necessary information covered in class?~~ Do the questions give students a chance to show what they know? Which questions are most difficult and why?

Yet another form of feedback is illustrated in Mina Shaughnessy's work on basic writing (1977). Mina Shaughnessy was not a researcher. She was at heart a teacher who made an enormous contribution to the teaching of remedial writing by systematically observing the types of errors students make. She concluded that:

"[remedial] students write the way they do, not because they are slow or non-verbal, indifferent to or incapable of academic excellence, but because they are beginners and must, like all beginners, learn by making mistakes. These they make aplenty and for such a variety of reasons that the inexperienced teacher is almost certain to see nothing but a chaos of error when he first encounters their papers. Yet a closer look will reveal very little that is random or 'illogical' in what they have written. And the keys to their development as writers often lie hidden in the very features of their writing that English teachers have been trained to brush aside with a marginal code letter or a scribbled injunction to 'Proof read!' Such strategies ram at the doors of their incompetence while the keys that would open them lie in view." (p.5)

Mina Shaughnessy used the data, collected from observing and talking with students, to make a major contribution to the teaching of basic writing. Most classroom research will not result in a book or even an article, but it should result in a better understanding of student learning by teachers who become curious about how students learn.

Moving up the scale of design sophistication in classroom assessment, we might arrive at an evaluation of a teaching experiment. An example is the study done by a group of math teachers who wanted to know if journal writing assignments improved learning in analytic geometry and calculus (Selfe, et. al., 1986). Quite importantly, these teachers also wanted to know how learning happened and whether the journal writing had any effect on students' attitudes. Did students feel better or more confident about writing tasks in general or about the subject matter they discussed in their journals?

The teachers designed an experiment with the appropriate control groups, asking the experimental group to write journal entries about their understanding of math concepts. The hypotheses were that writing about how to solve problems would help students think through the procedures and would result in better test scores, greater confidence, and more positive attitudes.

They found no statistically significant differences. Experimental students did not make higher test scores, did not show more favorable attitudes toward journal writing as a learning tool, and did not score significantly lower on a test of writing apprehension. Does the fact that traditional statistical

tests showed no significant differences mean that the experiment was a failure? Unfortunately it often means that the experiment never gets written up or published, but it does not mean that the results are not informative. In this case, qualitative analysis of the journal entries proved a rich source of data on how students go about solving problems.

"Because students had to record their responses in prose," wrote the teachers, "they could not hedge or guess at the correct answer or approach. Their level of understanding was immediately evident to the teacher. For this reason, journals were preferable to quizzes for the teacher. Journals also provided a record of individual learning styles, a written account of a student's progress as seen not only by the instructor, but by the student as well."

(Selfe, et. al., 1986, p.201)

The qualitative analysis of the data showed that students were engaging in some important kinds of learning not measured by standard quantitative methods. Teachers reported that students were:

1. Composing definitions of mathematical terms in a language that was their own.
2. Demonstrating understanding of abstract concepts through the use of analogies.
3. Evaluating the usefulness of the material in terms of their own experience.
4. Devising useful strategies for approaching math problems.
5. Using writing to think through solutions to problems.

(Selfe, et. al., 1986, p. 198)

Having read their report, I would like to do some research on those teachers. My hypotheses would be that they became more sensitive and aware of how students learned, that they enjoyed working on the project with their colleagues, that they became more interested in teaching and better teachers as a result of their classroom assessment.

Another example of classroom assessment is one carried out by an instructor in astronomy. In this study, the teacher is trying to determine which skills contribute to student success in her introductory class in astronomy. The study is a typical comparative design. The question for investigation is, What are the differences between students who receive As and those who receive Ds in the course? Data are collected via interviews, tests of visual/spatial ability and logical reasoning, questionnaires regarding skills and attitudes, student self-perceptions of ability to learn from the teaching strategies of the instructor, and observation of the class by a colleague. Ultimately, the goals are to determine the cognitive entry skills required for success in the course and to redesign the course and teaching strategies to reach more students. # An obvious extension of this design is for faculty members teaching sequential courses to work collaboratively on the identification of skills and attitudes needed for success in second-level courses and the appropriate preparation of students in first-level courses.

← Although the purpose of classroom assessment is to help individual teachers find answers to the questions that interest them, and its style is to be non-threatening and as private as the teacher wishes, there are many collaborative and cooperative projects that are possible.

Most assessment designs relate outcome measures to what the Teacher is trying to accomplish. Because college Teachers are rarely clear about their goals, beyond mastery of SM, my colleagues and I

~~As a part of our work on classroom assessment, we have~~

devised a Teaching Goals Inventory, consisting of 48 teaching goals gleaned from the literature. Examples of some typical goals are:

is trying to accomplish. Because college Teachers are rarely clear about their goals, beyond mastery of SM, my colleagues and I

←

My course will help students to:

1. Master knowledge in a discipline.
2. Improve powers of observation.
3. Develop the motivation and discipline for lifelong learning.
4. Acquire knowledge and skills needed for a job.
5. Develop effective study habits.

While I would expect teacher goals profiles on the TGI to differ from department to department, I think it would be interesting to see if departmental profiles "add up" to institutional goals, or if there are some goals promised or aspired to by the institution, for which no department is assuming responsibility. I would also expect the TGI to serve as a useful launching pad for departmental discussions on teaching goals and teaching improvement. Some departments may wish to carry out relatively complex evaluation programs to see if they are accomplishing their goals. This would require some agreement among teaching faculty as to what the goals are, and how student performance should be measured. One way to determine departmental goals is to simply list goals in order of the priorities assigned by the summation of individual faculty responses to the TGI. A better way is for the faculty to take a careful look at the collective responses and to discuss whether individual priorities do indeed add up to departmental goals.

In conclusion, the types of classroom assessment, ^{and research} projects that can be carried out are limited only by the creativity and energy of the faculty. The rationale for classroom assessment as I have thought about it--primarily at the level of higher education--is based on the following assumptions:

1. That the most obvious route to the improvement of learning is the improvement of teaching.

2. That the intellectual challenge of teaching should be emphasized, and that professional growth is possible through systematic study of the impact of one's teaching on learning.

3. That teachers of the various disciplines contribute to institutional goals in different ways. The physics department has different learning goals for students than the history department, and it makes sense to assess learning that is appropriate to the subject matter of the discipline.

4. That most faculty development programs now in operation analyze teacher behaviors. Classroom assessment offers an interesting alternative to instructional improvement by focussing attention on student behaviors and reactions.

5. That teachers who have designed their own classroom assessment projects are likely to consider the data relevant and useful, and therefore to use the results in the improvement of instruction.

6. That there is nothing so esoteric about classroom assessment that anyone capable of teaching a class cannot devise and conduct good classroom assessment projects.

7. That classroom assessment will be strengthened by relating it to departmental and institution-wide assessment and vice versa.

8. That there is a very good chance that thousands of small changes made in classrooms throughout the country will add up to more real reform than sweeping policy changes made at state and national levels.

In conclusion, classroom assessment is meant to be a sturdy vehicle to help teachers understand their students. It is not easily abused or damaged by teachers who want to know. At the same time, classroom assessment is fragile and can be damaged and even destroyed by those who want to make it into something it was never intended to be.

My hope is that we can keep the basic purpose of classroom assessment clear. It is not to do research. It is not to dazzle, impress, obfuscate, not to publish or enhance resumes, not to search for truth or the general laws of learning. The sole purpose of classroom assessment is to engage teachers intellectually in the profession of teaching and to bring an experimental and inquisitive dimension to finding better ways to cause learning.

4/20/87

REFERENCES

- Berliner, David C. "The Half-Full Glass: A Review of Research on Teaching." From Philip L. Hosford (Ed.), Using What We Know About Teaching. Alexandria, Va: Association for Supervision & Curriculum Development, 1984, pp. 51-76.
- Bloom, Benjamin S. "The New Direction in Educational Research: Alterable Variables." Phi Delta Kappan, February 1980, pp. 382-385.
- Centra, John A. "Student Ratings of Instruction and Their Relationship to Student Learning." American Educational Research Journal, Winter 1977, 14(1), 17-14.
- Cohen, Peter. "A Validity of Student Ratings in Psychology Courses: A Research Synthesis." Teaching of Psychology, 1982, 9(2).
- Dewey, John. The Sources of a Science of Education. New York: Liveright, 1929.
- Education Commission of the States. "Transforming the State Role In Undergraduate Education." Denver, July 1986.
- Schuster, Jack and Bowen, Howard. "The Faculty at Risk." Change September/October, 1985.
- Seldin, Peter. "Faculty Surveying Policy and Practices." Change, April 1984, 29-33.
- Selfe, C.L., Petersen, B.T., and Nahrgang, C.L. "Journal Writing in Mathematics" in Writing Across the Disciplines: Research into Practice. Young, Art and Fulwiler, Tuby (Eds). Upper Montclair, N.J.: Boynton/Cook, 1986.
- Shaughnessy, Mina P. Errors and Expectations. N.Y.: Oxford University Press, 1977.

4/20/87