DOCUMENT RESUME

ED 059 716 JC 720 071

AUTHOR Tolle, Donald J.
TITLE Evaluation: Who Needs It?
PUB DATE Sep 70
NOTE 15p.; Paper presented at a faculty workshop held at Mineral Area College, Flat River, Missouri, September 3, 1970

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Evaluation; *Evaluation Criteria; Evaluation Methods; *Evaluation Techniques; Faculty Evaluation; *Junior Colleges; *Teacher Evaluation

ABSTRACT

This is a review of the literature surrounding the development of techniques for, and changes in professional attitudes toward, the evaluation of community college faculty in their instructional role. Discussion focuses on the purpose, criteria, and process of evaluation. In discussing the purposes of evaluation, the author deals with teacher rating systems, philosophical conflicts between evaluation purposes and methods, and educational services' accountability and quality. The evaluation criteria discussion pinpoints attributes, abilities, and competencies that contribute to good teaching. Several lists of criteria are given as possible guidelines. The evaluation process is considered in the light of who is to do it and by what means. Several series of outlined procedures are given for instituting an evaluation process. (AL)

EVALUATION: WHO NEEDS IT?

Although the evaluation of instruction has a long history in American education, few people like to do the evaluating, and even fewer enjoy having their teaching effectiveness thus appraised by their fellow man.

In exploring the topic, we are dealing with an area of importance that has no one answer and perhaps no satisfactory answer for all people and for every institution. Somebody else's approach may not work well for anybody else, and a serious mistake would be made in taking someone else's "canned" method and applying it to this or any other community college. The best answer is for each college to consider the plans of others but to evolve its own rationale and procedures in the light of its own unique purposes, personnel, and programs.

The charge given to me was to join with you in looking at the purposes and procedures of evaluation for the improvement of instruction: Who should be involved in it, what are the pitfalls to avoid, what are the advantages and the disadvantages of evaluation? In short, a very large order. But let us begin by reviewing briefly some of the history of the subject.

Much research has been done in the past fifty years or so (most of it in the last twenty) to try to determine what makes a good teacher and how to evaluate teacher effectiveness. But little agreement has been reached except perhaps as to the characteristics of "good" teachers

---

Presentation at faculty workshop, Mineral Area (Community) College, Flat River, Missouri, on September 3, 1970, by Donald J. Tolle, Associate Professor of Higher Education, Southern Illinois University at Carbondale (Note: This presentation was designed to stimulate discussion on the part of the faculty.)

versus those of "bad" teachers--and, even there, there has been considerable disagreement as to what should be evaluated. Should teacher characteristics be appraised? Should teacher performance be evaluated? Should student learning be measured as a criterion of teacher effectiveness? Should some combination of these approaches be attempted? For what purposes should instructional evaluation be carried out: For improvement of instruction? For merit pay purposes? For advancement in rank and/or salary? For retention on the faculty? For self-improvement of faculty? For tenure or continuing contract purposes?

In one way or another teacher evaluation has been going on as long as we have had schools. Early supervision of instruction grew out of the responsibility of school boards to provide quality education; and such supervision was delegated to school administrators. The general tendency was for supervision to be authoritarian in nature and for teachers to be "inspected" rather than supervised in the best sense of that term. Too generally, the belief was that there was only one right way to teach--recitation, drill, memorization, etc. Creativity and differentiation of approach were exceptions rather than the rule. And only in the past half-century or so has there been a fairly general acceptance of the fact that not all good teaching is accomplished by the same means.

Cohen and Brawer in their ERIC monograph for the AAJC, entitled Measuring Faculty Performance, list 128 representative studies that have been carried out from 1915 to 1969, in regard to various facets of the topic. A background study of the evaluation of community college instruction, carried out by Highland Community College (Freeport, Illinois) this year, drew from thirty studies reported between 1928 and 1969, as well

as from its own survey of Illinois junior colleges. The following remarks will refer to findings reported in several of these studies, and it will be easy to note that the results are frequently disparate, nebulous, contradictory, and confusing--but possibly helpful. Generally speaking, the research and/or writing has dealt with three aspects of evaluation: Purposes or reasons for having it, standards or criteria, and procedures or processes or methods of evaluation.

Purposes

In 1950 the ASCD expressed its concern over the fact that teacher rating systems too often fail to allow for individuality and therefore may tend to exphasize conformity.[1*] The ACE in 1951 stated that "A semblance of an all-inclusive mechanistic concept tending toward standardization of approved practices, regimentation of teachers, and the suppression of originality can do great harm." However, the ACE also said that "most teachers can gain from the reasoned judgements of others concerning observable aspects of their work" and that for most teachers "evaluation should be a means of improving their effectiveness, thus lending to better security as well as personal satisfaction."[2]

The NEA, in 1969, took the position that using "subjective methods of evaluating professional performance" to determine salaries has a "deleterious effect on the educational process." But the NEA accepts tha fact that the teaching profession has as a major responsibility the evaluation of the quality of its services.[3] The fact of evaluation appears not to bother teachers as much as the question of how the results are to be used. If the purpose is simply the improvement of instruction,

*References cited are listed on page 15.

perhaps most teachers would agree to its value. However, when salary or merit pay or tenure or rank is involved, an attitude of distrust and insecurity frequently may exist. Cohen and Brawer report that the acceptance or rejection of the methods of evaluation often relates to the degree of acceptance or rejection of the purposes of evaluation.[4] Paul Dressel (1964) said, "Properly conceived, evaluation is not separate from instruction and learning; it is an integral part of both."[5] According to Barr, et al (1961), some form of evaluation is always in process, "whether made openly and carefully or made subversively and haphazardly."[6] Perhaps most would accept the fact that, if the primary purpose of evaluation is the improvement of instruction, it is a necessary and acceptable contribution to the good of the institution and of the profession. It is obvious, however, that, in some way or another, decisions must be made regarding placement on the salary schedule, retention and tenure, and merit pay and rank (when applicable). But it is generally at the point that evaluation is used for purposes other than improvement of instruction that the main problems arise. That fact, however, should leave this faculty in good condition, since your purpose is the use of evaluation for the improvement of instruction. The problem is that needed improvement of an individual's instruction frequently (and properly) may relate to any one or all of the other purposes just named--and administrators can't escape the responsibility of decision-making in these matters even though there may be a sharing and/or delegation of the authority involved.

One necessary purpose of evaluation may well come increasingly into the picture during the next several years, and that is the concept of

accountability for educational services and the quality of their outcomes. Taxpayers and legislators are becoming restive over the high cost of education at all levels, the unrest and disruptions taking place on campuses all over the country, and the strong criticisms of lack of relevance being directed against a great many educational programs and teaching methods. Although the junior colleges have thus far been only lightly involved in such attacks, there is no room for smugness or complacency in any educational institution in the country.

The Council of North Central Junior Colleges which meets in Detroit next month has taken as its theme: "Accountability: Rights, Responsibilities, and Relevance --- the 3 R's for the 70's." You may be interested in some of the topics that will be discussed at the conference:

"Accountability for What?"
"Collective Bargaining and Accountability."
"Faculty Accountability in the Community Colleges."
"The Dean of Instruction and Accountability for Instructional Outcomes."
"Student Rights, Responsibilities, and Accountability in the Face of Disruptions."
"Institutional Accountability in Relation to Minority Groups."
"Presidential Accountability in the Face of Increasing Pressures from Board, Faculty, Student, and Taxpayer Groups."

Criteria for Evaluation

Many studies have attempted to pinpoint those characteristics, attributes, abilities, and competencies that contribute to good teaching. Although there is no consensus, perhaps three general categories of criteria should be considered in instructional evaluation (as given by Bannister, 1961):

1. Classroom atmosphere - a "climate" conducive to student ease, where students feel they have the respect of their instructor and classmates, where they are challenged

by their work, where they are confident they can succeed, and where they experience gratifying success.

2. Instructor - a person who is tolerant, reasonable, approachable, who possesses mastery of field and understanding interest and enthusiasm for the subject, who is thoroughly prepared for each class, and who conducts each class efficiently without annoyances or mannerisms which divert attention.

3. Course - one which has clearly defined objectives and standards which must be attained, which utilizes methods and material adapted to specific needs of the student but allows for individual differences, in which there is student participation, reviews at regular intervals, fair tests returned promptly, in which the interrelatedness of knowledge and relation to daily life are stressed, and in which students are appraised periodically of the quality of their progress.[7]

In 1961, Barr gave a comprehensive summary of a large number of studies and concluded that:

1. Judgements as to whether a teacher is effective depend upon the criteria used.

2. Developing evaluative criteria is a difficult and complex undertaking.

3. There is much unevenness in the abilities of teachers--they may be low in some, high in others.

4. Teacher acts have an appropriate aspect not good or bad in general but in relation to purposes, persons, and situations.

5. There may be enough individuality in raters to produce differences in the ways in which teachers succeed and fail.

6. Since teaching does not take place in a vacuum, other factors beside the teacher may influence teaching and learning.[8]

With all the research that has been done, every study still seems to end with a recommendation that even more work be carried out to identify the criteria that should be used in the evaluation of instructional effectiveness.

Fritschel (1967), in trying to describe minimum standards of competence, noted that such definition can not be on the basis of what a

teacher is but by what he does, what action he performs, what role he plays, and how he carries out his responsibility. He defined these areas of minimum competence as reflecting a person who (a) is a director of learning, knowing about his learner and how he learns, knowing his subject matter, and being a member of a teaching team, (b) has human relations skills, and (c) is an agent of change.[9]

A great many evaluative instruments have been devised to assess the quality of instruction. According to Kent (1967), the most common areas covered have been course goals, content, materials, assignments, instructor behavior, mastery of subject, personal traits, and relationships with students.[10] Kent also gave attention to the "unauthorized" student ratings which increasingly are being published on college campuses with the avowed purpose of improving teaching and of aiding students in their selection of instructors. Such ratings contain comments about individual teacher's effectiveness, and in most cases they include information about the instructor's enthusiasm for his subject, his organization, his manner of presentation, his fairness in grading, and his personal traits (with "warmth and friendliness" being particularly valued).

An extensive study carried out at the University of Toledo attempted to identify those behaviors which reflect positive teaching and to determine their relative importance. Over 13,000 such "behaviors" were reported by students, alumni, and faculty, and then grouped into sixty criterion statements, which are rated as to their relative importance to good instruction.[11]

Highland Community College (Illinois) tested the Toledo listing with its faculty, students, and administrators and found that a high correlation

existed between the ratings of the teaching behaviors by the Toledo and the Highland academic communities. The first 15 criteria in rank order (according to the Highland survey) included every item that appears commonly in many other evaluative instruments, indicating that it is quite possible for a college to develop a list of criteria which can identify strengths and weaknesses in the instructional process. The first fifteen items in rank order were:

1. Being well prepared for class
2. Treating students with respect
3. Acknowledging all questions to the best of his ability
4. Using teaching methods which enable students to achieve objectives of the course
5. Being fair and reasonable to students in evaluation procedures
6. Demonstrating comprehensive knowledge of his subject
7. Establishing sincere interest in the subject being taught
8. Being readily available for consultation with students
9. Constructing tests which search for understanding on the part of the students rather than rote memory ability
10. Encouraging intelligent independent thought by students
11. Patiently assisting students with their problems
12. Communicating effectively at levels appropriate to the preparedness of students
13. Motivating students to do their best
14. Organizing the course in logical fashion
15. Accepting justified constructive criticism by qualified persons[12]

But can the assessment of instructional effectiveness be limited to such characteristics as these? Cohen and Brawer (in Measuring Faculty

Performance) say, "No." They say that teaching can be defined realistically only as "causing learning," and they avow that student gain toward specific learning objectives is the only really meaningful criterion that can be used in appraising the effectiveness of teachers.[13] They maintain that such learning can be measured in objective fashion. Partly from this school of thought has come the great emphasis in recent years upon the so-called "behavioral" or "performance" or "measurable" objectives in education.

Certainly a renewed focus upon objectives, especially in measurable terms, has merit for each teacher in the planning and conduct of his courses. Such an approach requires the teacher to think clearly about purposes, content, means, and outcomes; and to plan carefully is a continuing need for all of us who teach.

But, in my view, the question is whether we have to "swallow whole" this one approach in either our teaching or our evaluation of teaching effectiveness. It is quite possible that the intangibles of the teaching-learning outcomes are highly important, yet not susceptible to objective measurement. Each institution will simply have to decide for itself what criteria are to be used for the evaluation of instruction, but I suggest that a multi-directional approach may bear best fruit.

The Process of Evaluation

After decision has been made as to the purposes for which evaluation is to be carried out and as to the criteria which are to be used, the difficult remaining task is development of the process by which evaluation is to take place. Who is to do it and by what means?

In 1960, AACTE reported the results of a survey it had taken to find

out how colleges sought out the information used in appraising teacher ability for promotion and pay purpose. Over three-fifths reported that such information was received by "grapevine," with administrative observations, student achievement, and student rating being less used sources of data.[14]

An ACE survey in 1966 revealed that college undergraduate instruction is evaluated in an equal number of instances (80%) by the department chairman and by the dean. Colleague evaluation, publication, and informal student opinion each ranked at about the 40th percentile. In the case of junior colleges, the dean evaluated in 82% of those reporting, with department chairmen participating in 65% and colleague evaluation taking place in less than 30% of the total. Publication was an insignificant factor in junior college evaluation, but class visitation was used by 42% (compared to only 12% in all colleges, both junior and senior, participating in the survey).

In the Highland Community College's survey of Illinois junior colleges, of the 24 which responded, five had not yet developed a system of evaluation, seventeen use some form of evaluation by administrators, and six use evaluation by colleagues, eight by students, and six by self-appraisal. Ten of the 19 which have an evaluation system allow or require class visitation by administrators, usually the adacemic dean or department chairman.[16]

The reasons for which evaluation is carried out frequently result in varying types of assessment techniques for different purposes. For instance, those purposes that don't relate to salary often emphasize self-evaluation and teacher-administrator cooperative plans. Those that have to do

with merit pay usually involve some sort of ranking of teachers and use of evaluative boards or committees.

To be effective, evaluation for the improvement of instruction must be a continuing process--not a fitful or sporadic one. It is this fact which causes some of the teacher criticism of classroom visitation by administrators, especially if it is done infrequently and without pre- and post-conferences. Fawcett (1964) said, "Hasty evaluation of performance during performance is not complete, fair or accurate. There should be a long sequence of evaluation."[17] This requirement is hard to meet because of the many other demands upon administrators' time, but, unless class visitation can be conducted in this manner, it will remain one of the poorer evaluative techniques.

On the other hand, this very emphasis upon the need for observation over a long period of time buttresses the case for student evaluation of teacher effectiveness. Common arguments against student ratings (as given by Kent) are that (a) Students are not able enough nor mature enough to judge effectively; (b) Students have no right to evaluate instructors; (c) Rating instruments are subject to bias and unreliability. However, several studies have found that students can evaluate fairly and with reliable results (e.g., Casey, 1966; Henderson and Chambers, 1951; Kent, 1967).[18]

Harold Howe, former Commissioner of Education, backed the right of students to evaluate instruction when he said, "The opinion of those who eat the pudding certainly ought to be considered if we want to know how it tastes."[19] As one who has for many years used student evaluation of my own teaching effectiveness, I can testify to its worth. We seldom see ourselves as our students view us every day--we normally think we're better than we

are, and a periodic, anonymous rating by our students can help restore perspective and, perhaps, some measure of humility. Student criticisms and suggestions for improvement can be truly helpful to most of us if we will simply tap this source of instructional improvement.

Certainly, students are at least as subject to biased judgments as are other evaluators. We all see the same thing from different vantage points and with our perceptions colored by our values and experience--in spite of conscious effort to be "objective." It is important, therefore, from both the institutional and the individual standpoint, that evaluation of instructional effectiveness derive from more than one source so that various facets and dimensions of the learning-teaching situation may have better chance of fair assessment.

The Highland Community College report, in citing a study by Herge (1965), had this intriguing footnote: "In all honesty. . .one must report a recent study which indicated that students in classes of teachers receiving high student effectiveness rating did not necessarily learn more than those in classes with teachers rated ineffective. Does this suggest that students often learn in spite of the quality of instruction they receive?"[20]

An example of instructor self-evaluation, reported by Anderson in 1964, has interesting possibilities for any teacher's use. In this experiment, each instructor rated himself on a 7-point scale for the following attributes: speaking voice; mannerisms or pleonasms; knowledge of subject matter; personal enthusiasm; enthusiasm engendered in students; digressions; handling of questions; and general atmosphere created in the classroom. The instructor then make audio-tapes of two 1-hour class periods. After listening to the tapes, he completed another rating sheet and compared the two

ratings. Although no statistically significant differences were found between the "before" and "after" ratings, more than half the faculty appeared sensitive to the information obtained from the tapes. Of the 19 instructors involved, five rated themselves more favorably the second time, six rated themselves less favorably, and eight did not change their ratings. The instructors concluded that the exercise was of value to them. Anderson lists the advantages of this technique as follows:

1. evidencing interest in the teaching process itself by the adminsitration,
2. indicating confidence by the administration in the faculty's ability to evaluate themselves as professionals and make self-indicated improvements,
3. giving the faculty a workable and frequently interesting method whereby they may improve themselves,
4. preservation of anonymity by faculty, thus forestalling feelings of "big brother" watching,
5. establishing essentially a self-operating and perpetuating system not calling for a great amount of time,
6. placing of the dean in the position of being called in for aid by a motivated faculty member, rather than being looked upon as an intruder with unwanted advice, and
7. providing specific and concrete examples (preserved on tape) of problems which can be referred to on replay, without having to rely on notes or faulty memory.[21]

This technique, with additional experimentation (preferably using videotape), could be a valuable tool in producing increasingly better instructors.

We have been looking at various facets of evaluation for the improvement of instruction, during the past several minutes--more minutes, perhaps, than you care to count at this point. So let me hasten to a conclusion of this formal part of the presentation by suggesting to you some "do's" and "don't's" of evaluation.

Don't be afraid of the idea of evaluation--embrace it (at least gingerly if not with passion).

Don't take over someones else's plan--develop your own.

Don't turn over to a small committee the complete responsibility of determining your evaluation system. A steering committee to coordinate the development of the plan would be good; but work toward general involvement of faculty and administration in the planning.

Don't depend on just one type of evaluation of instruction, although it is sometimes better to start small and build gradually. Give consideration to several possibilities: student evaluation, self-evaluation, peer evaluation, administrative evaluation.

Don't force a system of evaluation on the faculty, but seek prior, general acceptance. Nothing has less meaning or does more harm than appraisal practices which are distrusted by those being appraised.

Don't expect perfection of your plan; settle for a little less. On the other hand, don't settle for as much less as the maiden lady I heard about one time. She loved a rather shiftless, ne'erdowell sort of fellow. He was nice enough, but he could never seem to get and keep any money. So the good lady decided she was going to have to take action if he was ever going to be able to marry him, and she said, "John, if you don't go away, get a job, and save $3,000, I'm not going to marry you." He said O. K. he'd try, and he left town. By the time he'd been gone a couple of months, the lady missed him terribly and thought maybe she'd made a mistake. Then one day she saw him walking down the road and ran out joyfully to greet him. She said, "Oh, John, have you saved the $3,000?" He said, "No, but I have $38." She said, "That's close enough." . . . . . .So don't settle for <u>too</u> little in your evaluation scheme.

Now I'd like to open the floor for discussion of any of the points I've raised.

1. Reported in "Evaluation of Community College Instruction: A Background Study," Highland Community College (Freeport, Illinois), 1970, p. 2.

2. Ibid.

3. Ibid.

4. Cohen, Arthur M. and Florence B. Brawer, Measuring Faculty Performance, American Association of Junior Colleges (Washington, D.C.), 1969, p. VIII.

5. Reported in Highland Community College, op.cit., p. 2.

6. Ibid.

7. Reported in Junior College Research Review, January 1970, p. 1.

8. Reported in Highland Community College, op.cit., p. 3.

9. Ibid., p. 4.

10. Ibid.

11. Ibid., p. 5.

12. Ibid., p. 20.

13. Cohen and Brawer, op.cit., p. 51.

14. Reported in Highland Community College, op.cit., p.6.

15. Ibid.

16. Ibid.

17. Ibid., p. 8.

18. Ibid.

19. Ibid.

20. Ibid.

21. Junior College Research Review, op. cit., p. 2.

bdg