DOCUMENT RESUME

ED 059 708 JC 720 061

TITLE	Guidelines for Evaluation in California Community Colleges.
INSTITUTION	Los Angeles Pierce Coll., Woodland Hills, Calif.
PUB DATE	71
NOTE	18p.; A report prepared for the Los Angeles Community College District personnel: faculty, administration, and board of trustees, 1971-1972
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	*Evaluation; *Evaluation Techniques; *Evaluative Thinking; *Junior Colleges; *Teacher Evaluation

ABSTRACT

This report was prepared in part as a response to a State of California legislative mandate calling for evaluation in community colleges. The preparation of evaluative guidelines was the primary goal and that effort was based on four action steps: (1) ascertaining faculty and administrative attitudes locally towards evaluation; (2) performing a review of evaluation literature; (3) surveying faculty and administrators in all California community colleges on their attitudes regarding evaluation; and (4) drawing conclusions from all the data. These actions resulted in the development of guidelines relating to: principles of evaluation, purposes of evaluation, procedures of evaluation, and criteria of evaluation. A preliminary draft of a code of ethics is appended to the main body of the report. (AL)

Guidelines for Evaluation

in

California Community Colleges

A Report Prepared for the
Los Angeles Community College District
Personnel: Faculty, Administration, and
Board of Trustees
1971-1972

Prepared by:

Innovations Committee
Los Angeles Pierce College
Woodland Hills, California

Mary Baump, Nursing Department
Jill Bohlander, English Department
Frank Chookolingo, Political Science Department
Orolyn Clark, Secretarial Science Department
Ronald Farrar, Foreign Language Department
Eugene Hinkston, Political Science Department
Elaine Lopeirs, History Department
James Trudell, English Department

Benson R. Schulman, English Department and
Chairman of Committee

EVALUATION IN CALIFORNIA COMMUNITY COLLEGES

Many of us begin study of evaluation with a feeling of regret, an attitude that we have more important things to do than respond to legislative imposition. Perhaps that is how the Pierce College Innovations Committee and, later, The Los Angeles Community College District Academic Senate initiated study of evaluation--as a response to an imposition, as it were. No matter! We now face the mandate of evaluation and, hopefully, we shall act wisely.

As we think back, these introductory comments, and many others, were included within the Pierce Innovations Committee discussions during spring 1971 when we first discussed evaluation of instruction. We concluded that respectable study might enable us to contribute to the advancement of our profession, and to assure that mandated evaluation might become a respectable procedure as opposed to an instrument of oppression. We decided, therefore, to secure a broad base of data as follows:

1. Ascertaining faculty and administration attitudes toward purposes, procedures and criteria for evaluation through a pilot study by questionnaire among the eight colleges in the Los Angeles Community College District,
2. Performing a review of literature about evaluation,
3. Constructing a final questionnaire based upon the pilot study for distribution to all community colleges in California, and
4. Summarizing and concluding from all data in order to recommend guidelines for evaluation which are based upon a reasonable foundation.

A few explanatory comments about the conditions of the study may be worthwhile at this time. First, it was conducted by an all-instructor committee, without any administrative control whatsoever, but with the encouragement and support of John Nicklin, President of Pierce College, and Donald Click, Chancellor of the Los Angeles Community College District. Second, one of the participants, James Trudell, is performing a doctoral dissertation on our study at the University of Southern California, and he accepted major responsibility for surveying the literature on evaluation. Third, the study is not quite completely finished since

returns from the state-wide questionnaire are still arriving, but our data are based upon responses from more than fifty per cent of all of our community colleges.

The first step in the study was to prepare a pilot questionnaire for distribution within the Los Angeles Community College District during summer 1971. The instrument contained fifty questions, some repetitive to provide a checking feature. Responses were solicited in three categories from faculty and administration to determine attitudes 1) toward evaluation generally, 2) toward methods and purposes of evaluation, and 3) toward criteria for evaluation. The questions, it should be added, were mixed in random fashion to avoid stimulating responses. Finally, a total of 164 out of a possible 256 responses was tabulated, including the following categories of personnel: 16 administrators, 42 department chairmen, 68 tenured faculty, and 38 probationary (including one substitute) faculty.

The findings from the pilot study are briefly summarized in several categories. The first concerns attitudes toward evaluation generally:

1. Too frequent evaluation interferes with academic freedom, adversely affects morale, and resembles surveillance. (faculty and administration disagreed)
2. Evaluation procedures should avoid irrelevant political, social, personal views and habits of the individual. (all categories of respondents agreed)
3. Inefficient or too frequent evaluation of tenured instructors will pose a heavy burden on department chairmen and administrators by demanding additional duties and increased paper work. (a comment freely added by many respondents)
4. Frequent evaluation can stifle creativity and innovation among teachers.
5. Unorthodox instruction is harder to evaluate than conventional. (instructors and administrators disagreed)
6. Most existing evaluation practices exist for administrative convenience.
7. Self-evaluation should be included in the total process. (volunteered by many respondents)
8. Instructors do not do their best work when being observed. (instructors disagreed with department chairmen and administrators)

Responses concerning purposes of evaluation showed greatest agreement among all categories of respondents. These are identified in rank order:

1. improvement of instruction
2. judgment of subject matter preparation
3. retention in employment
4. judgment of methods of instruction

Other possible purposes of evaluation were largely rejected!

Evaluation procedures identified, involved the following, listed in order of preference:

1. peer instructors
2. department chairmen
3. administrators
4. students (preferred by over 50 per cent of probationary instructors)
5. self-evaluation

Methods of evaluation included:

1. classroom visitation
2. assessment of instructional success through judging student attainment of objectives and criteria of instruction
3. conferences between instructor and his evaluators

Criteria of evaluation, approved nearly unanimously by all respondents, included:

1. ability to relate to students in class
2. knowledge of subject matter
3. effectiveness of teaching methods
4. students' performance in relation to course objectives

A few final comments on the results of the pilot study may be useful here. There was general agreement that the same evaluation criteria should apply to tenured and non-tenured instructors. However, there was no agreement on whether criteria should be the same for vocational and non-vocational instructors. Frequency of evaluation was a matter of concern too: most respondents felt that probationary faculty should be reviewed each semester, but department chairmen and administrators favored yearly review of tenured faculty while tenured faculty preferred intervals of two years or longer.

The pilot study within the Los Angeles Community College District revealed

a number of attitudes and preferences which are not, perhaps, surprising to most educators; but what did we obtain from our review of literature on the subject? May we present some samples?

> Sprinthall, Whiteley and Mosher called the literature on teacher effectiveness almost unmanageably extensive, since more than two thousand studies (Ours, presumably, is 2001!) have been made since 1900. However, they found little research which could be called conclusive in determining exactly what effective teaching is and how it can be reliably measured and predicted. (Furthermore) Aside from finding that students learn better when teachers are kind, cheerful, and sympathetic, researchers know little else. (Finally) The greatest number of studies has been concerned with the collection of subjective opinions from experts, laymen and students.

The highest-rated teachers show several typical characteristics: good instruction, good classroom management, professional attitude, wise choice of subject matter, good personal habits, ability to cooperate, and good health. Furthermore, in the words of Barr, "Teaching is an exceedingly human task."

Ryans identified highly-rated teachers as follows:

1. They are favorable in opinions of students.
2. They favor democratic classroom procedures.
3. They show high mean verbal intelligence scores.
4. They show superior emotional adjustment.

Methods of evaluating have been the subject of much speculation. Classroom observation, for example, perhaps the most common means used, was often challenged as unreliable and insufficiently analytical. Furthermore, Ryans indicated that even when highly trained observers were used, they frequently showed wide discrepancies in their reports about the same individuals.

Anderson criticized evaluation criteria because they depend upon subjective, personal judgment and because there is disagreement about "which traits, behaviors, and qualities are essential to success in teaching and the extent to which each contributes." In seeking better, more precise ways, some advocate measuring effectiveness of instruction in relation to student performance, the "objectives and criteria" approach. Anderson, however, attacked this method too, suggesting that students are not uniform in ability and, therefore, cannot

be expected to perform equally no matter how excellent (or poor) the instruction. (Does this resemble the community college population problem?) Finally, Gustad noted that although the majority of collegiate institutions placed great weight on teaching ability, they had no effective method of evaluating it.

In a study of rating practices in 584 colleges, Gustad reported that fifty-two per cent used classroom visits. Other methods included 1) colleagues' opinions, 20 per cent; 2) course syllabi and examinations, 32 per cent; 3) grade distributions, 20 per cent; and 4) student ratings and opinions, 15 per cent. In the face of the reported 52 per cent emphasis upon visits to classrooms, should we be dismayed by Cohen and Brawer's statement that rater bias causes a serious problem where teacher evaluation is concerned? On the other hand, Cohen and Brawer assert that raters should concentrate on the products of learning and teaching rather than on any other means of evaluation; but how does this jibe with opinion of Anderson, reported earlier, that this method is poor because students are not uniform in ability?

What should the purposes of evaluation be? In a study conducted during 1955-56, Davis found that teacher ratings were used for the following reasons:

1. supervisory aid..... 77%
2. placement of non-tenured teachers..... 71%
3. promoting probationary teachers to tenured status.. 56%
4. promoting teacher..... 36%
5. determining pay increments..... 16%
6. selecting those who would be given super-maximum salaries..... 07%

Nowhere was improvement of instruction even mentioned! Furthermore, Wisgoski indicated that evaluation should improve teaching, inspire mutual trust, be institution-wide, and should include all members of the staff and administration.

In further pursuit of the purposes of evaluation, we find that a 1915 resolution of the NEA stipulated that ratings should not, "Unnecessarily disturb the teacher's peace.", but that they could legitimately be used for setting salaries and granting tenure. On the other hand, we find the following included

in a report to the American Association of Junior Colleges as recommended purposes:

1. to determine achievement of the objectives held by the school,
2. to provide the basis for rewarding superior service,
3. to provide the basis for self-improvement,
4. to provide the basis for motivation,
5. to provide the basis for inservice and supervisory activities,
6. to provide the basis for administrative decision, and
7. to provide the basis for judgments

What should the process of evaluation include? Wisgoski suggested self-evaluation, student evaluation, and classroom visits by qualified and unbiased observers. Gage recommended correlating student rating and student achievements. Shawl held that evaluation can be performed best by determining what has happened to the learner. McNeil suggested a plan whereby the instructor and his evaluator would agree in advance about what would show that the instructor has effected positive changes in his students. On the other hand, Dressel cautioned that any evaluation plan faces a problem: "The danger is that the college teacher does what he sees is rewarded." A further caution resulted from a Yale study in 1965, in the form of a conclusion about student evaluation: it should not be used as a basis for tenure and promotion because students lack expert knowledge and may, however unwittingly, cause harm. What, from these samples, may be concluded about the process of evaluation?

Evaluation criteria have been delineated in the literature in various ways. Barr identified three categories of criteria: personal qualities, competency as a director of learning, and the results of teaching. In a questionnaire distributed by Astin and Lee, we find the following reports from colleges: 98.2% of respondents chose classroom effectiveness, 69.2% named personal attributes, 42% named student advising and counseling, and 41.5% preferred campus committee work.

Drayer reported somewhat differently from a study of instructor qualities conducted among 148 college students:

1. 105 appreciated "presenting material effectively" first,
2. 54 identified "good command of subject,"
3. 53, "gave personal help to students,"
4. 41, "tests were fair,"

5. 33, "made the course interesting," and
6. 29, "impersonal, yet friendly."

In another study, Gough concluded that the prediction of performance of teachers through psychological inventories is an unsolved problem. Finally, Krupka surveyed instructors and students concerning criteria, showing much similarity of responses:

	faculty ranking	student ranking
1. knowledge of subject	1	1
2. ability to arouse interest	2	2
3. classroom presentation	3	4
4. organization of course	4	3
5. relevance of assignments	5	unranked
6. willingness to help	6	5

What additional ideas from the literature may help us? Biddle, Ort and others reported that the best evaluative technique is observation by supervising teachers. Astin and Lee, on the other hand, stressed student evaluation as based upon the greatest opportunity to observe the instructor. Eble reported that well-formulated and administered student evaluations are useful to improve teaching. In other places, one finds emphasis upon student performance, without any classroom observation at all being recommended. However, in general, the review of literature shows that evaluation is not very scientific or precise. If, as so many have said, teaching is an art not a science, then, perhaps, really objective attempts to measure it are not possible with present knowledge.

The last step in our study consisted of mailing questionnaires to the 94 community colleges in California. Ten were sent to each college as follows:

1. one copy to an administrator with instructional responsibility
2. one copy to a department chairman
3. three copies to probationary and substitute instructors
4. five copies to tenured instructors

As of this writing, 502 of a possible 940 questionnairs have been returned and analyzed, a total of 53.4 per cent. Of these, 203 were from rural districts and 299 from urban districts. The precise breakdown is as follows:

Position of Respondent	Rural Districts	Urban Districts
Administrators	22	37
Department Chairmen	32	34
Probationary and Substitute	53	84
Tenured	96	144
Totals	203	299

We originally sought data about rural-versus-urban because we thought there might be differences between attitudes of people who worked in different areas, but, surprisingly or not, there were no differences attributable to areas. Therefore, we shall report our data on the basis of position.

In general, the results of the state-wide study tended to confirm the results of the pilot study conducted within the Los Angeles Community College District during the summer of 1971. A sampling is provided for review.

1. frequent evaluation resembles surveillance. (faculty and admin. agreed)*
2. evaluation procedures should avoid political tests (faculty and admin. agreed)
3. unorthodox instruction is harder to evaluate than conventional (faculty and admin. disagreed)
4. existing evaluation serves administrative convenience (faculty and admin. agreed)
5. self-evaluation should be included within the total process (faculty and admin. agreed)
6. instructors do not do their best work when being observed (a majority agreed with the exception of tenured instructors)

Responses to a request for ranking of purposes of evaluation showed substantial agreement:

1. improvement of instruction: 99%+ administrators; 85% department chairmen; 81% probationary and substitute instructors; and 86% tenured
2. judge subject preparation: ranked second or third by all groups
3. retain in employment: ranked last by all groups
4. determine instructional assignment: ranked second or third by all groups

*response differs from pilot study

All evaluation procedures identified were substantially approved by all groups, but percentages varied:

Procedure	Admin.	Dept. Chairm.	Prob.&Sub.	Tenured
classroom visitation	78%	79%	78%	77%
evidence student achievement	81.5%	59%	63%	69%
participation by peers	83%	71%	75%	82%
participation by students	76.5%	68%	73%	71%
self-evaluation	81.5%	82%	74%	76%

Should student evaluation be official or confidential to the instructor?

Group	Confidential	Official
administrators	41%	59%
department chairmen	52%	48%
probationary & substitute	51.5%	48.5%
tenured	49.5%	50.5%

Should peer evaluation be official or confidential to the instructor?

Group	Confidential	Official
administrators	19%	81%
department chairmen	28%	72%
probationary and substitute	37%	63%
tenured	32%	68%

How should methods of evaluation be selected?

Methods	Admin.	Dept. Chair.	Prob.&Sub.	Tenured
administration alone	0	1%	2%	3%
faculty alone	0	3%	8%	6%
faculty & administration	33%	23%	22%	23%
faculty, admin., & students	67%	73%	68%	68%

Although there was much agreement concerning use of students and peers, and methods of evaluation, when the criteria of evaluation uncovered in the pilot study and the literature were presented to the state-wide group for ranking, the figures were not conclusive. In general, the greatest uncertainty was about evaluation criteria. These are:

1. ability to relate to students in class: ranked second by 57% of admin.
2. student achievement based on instruction: ranked first by 84% admin. and first by 46% tenured instructors
3. methods of instruction: ranked fourth by 50% of admin., fourth by 45% dept. chairmen, fourth by 53% prob. & sub., and fourth by 50% tenured
4. knowledge of subject was ranked most inconclusively by all

A few general comments may sum up the reactions to the state-wide questionnaire. First, respondents showed a desire to broaden the base of participation by personnel. Second, they recognized that existing evaluation practices are not adequate. Is this because practices are too rigid and do not allow for appropriate individual variations? This question suggests that many respondents believe that methods and criteria ought to be eclectic, according to what applies best to individuals.

A final consideration relates to the matter of "conducting evaluation for administrative convenience. It was rejected by all respondents, administrators and faculty alike! Does this not "mandate" a severe restructuring of community college evaluation practices?

Conclusions and Recommendations

Drawing conclusions from data such as those we have assembled is fraught with difficulties. First, there was a clear lack of precise evidence in the literature or the surveys to support claims of proof of anything; that is, we cannot show unequivocal support for one position or procedure as opposed to others. Second, the process of evaluation is fallible because instructors and those who

would evaluate them are human and fall easily into error. Third, as educators ourselves, we are troubled because, although we "know" that a sound evaluation procedure can be beneficial, we instinctively abhor the idea that we must somehow measure up to what others think we should attain. We are, it should be added, fully aware of the differences in perception from which people initiate study of educational performance. In this context, we shall, nevertheless, enunciate certain principles, purposes, procedures and criteria for evaluation.

Principles of Evaluation

1. Educational evaluation is subjective, not precise, and should therefore include provisions for checks against abuse.

2. Evaluation procedures may seriously affect instructor and administrator morale and should, consequently, be adopted by broad consensus, not imposition.

3. No single evaluation procedure will enable all instructors to be evaluated justly or all evaluators to judge wisely.

4. Effective evaluation is complex and difficult: time and necessary financial support must be provided or the process will not function well.

5. Those who evaluate should have clear qualification for doing so and those being evaluated should have advance knowledge of all criteria which they are expected to meet.

6. Provision should be explicated for appeal against poor evaluations.

7. Poor evaluations, particularly, should include specific evidence to support conclusions and should be based upon a number of judgments of unbiased individuals.

8. Provision should be made for practical remedies such as in-service activities to assist those whose performances are judged ineffective.

9. All participants in evaluation should share equally in selecting criteria, including evaluators and those being evaluated.

10. Those being evaluated should have some options in selecting their evaluators and the procedures to be used.

11. Evaluation should be confidential.

Purposes of Evaluation

Purposes of evaluation should be acknowledged by consensus and should help to improve the conduct of the profession. They follow:

1. Improvement of instruction
2. Judgment of effectiveness of subject matter preparation
3. Judgment of effectiveness of methods of instruction
4. Judgment of whether or not to retain in employment

Procedures of Evaluation

Evaluation procedures should contribute to high morale by being fair, by allowing some choices by participants, by not placing undue pressures on participants, and by being conducted no more frequently than necessary. The principle is acknowledged that those who are judged relatively ineffective should be evaluated more frequently than those who are judged fully competent.

Procedures follow:

1. Classroom visitation, if utilized, should be made by judges most competent to determine effectiveness. Therefore, the use of departmental or divisional colleagues receives first emphasis, with administrative participation in a lesser role.

2. Evaluation without classroom visitation is legitimate and may be arranged for those instructors who wish to stipulate the objectives and criteria governing their instruction and who elect to be judged by the performance of their students.

3. Self-evaluation should become a regular practice by all instructors to stimulate thoughtful and regular self-appraisal. After completing a self-evaluation, each instructor should be prepared to discuss his findings with trusted colleagues and/or administrative personnel.

4. Student evaluation* should be used by all instructors because students are those most intimately exposed to instructor performance. However, because so much objection has been raised to possible pressures created by published student ratings and opinions, the procedure should be carefully controlled.

*The Pierce College faculty showed opposition to student evaluation.

Before engaging in evaluation, students should be instructed as to the nature of their task and cautioned against purely emotional judgments, pro or con. These evaluations should be collected and studied by the instructor for his own guidance. Finally, the instructor should be encouraged (but not required) to discuss the results with trusted colleagues and/or administrative personnel.

5. Discussion of observations completed and documents collected during the evaluation process should be the final step, with ample opportunity to consider all data before conclusions are drawn and additional steps, if any, are selected. The final discussion should include one or more trusted colleagues, the responsible administrator(s), and the person being evaluated. No entries should be included in an evaluation record of which the person being evaluated is not informed.

6. Evaluation records should be confidential, but they may be released with the permission of the person whose records they are. Records should be current, no longer than two years old if adverse in nature and the deficiencies earlier recorded were corrected.

Criteria of Evaluation

Criteria for evaluation are, perhaps, the most troublesome aspect of the entire problem. We have already reported from the literature that measuring instructional effectiveness is extremely difficult because there is little or no agreement about how to measure with objectivity and precision; nor, for that matter, is there solid evidence as to the attributes which make for good or bad instruction. Let us, therefore, make an admission: evaluation is subjective and if we really wish to make it workable, we must select criteria which can be utilized in as non-subjective a manner as possible.

We must, for example, strive to eliminate factors which have not been proven to influence instruction directly: personality, political and social beliefs, mode of dress, and similar personal qualities which even psychiatrists, sociologists and "you-name-it-ologists" of every color and persuasion differ about. We must, in effect, make the deliberate assumption that personal factors show varying effects among people and within the same person at different times; and, furthermore, similar personal qualities are often found in instructors whose effectiveness varies widely. We must, in short, reject the urge to evaluate what we cannot

measure, with one possible exception: when an instructor's personal qualities show clear violation of professional ethics or of the law. Without apology, consequently, our recommended criteria for evaluation of instruction are few and directly related to the educational process:

1. The instructor's ability to relate to his students,
2. The instructor's knowledge of his subject matter,
3. The effectiveness of the instructor's teaching methods, and
4. The response of the students to the instructor's performance.

One Final Recommendation: A Code of Ethics

This committee has made a number of comments urging major faculty responsibility for the evaluation process. Furthermore, since administration has a legal responsibility for the supervision of faculty, its role in evaluation is of great significance also. It seems important to the comittee, therefore, that both faculty and administration develop a code of ethics and, perhaps, a system of due process to delineate rights, responsibilities and procedures for the protection of all involved in evaluation; and for legal purposes the board of trustees of the college district must ratify the code and the due process procedure. If such a code and due process procedure are enacted to govern evaluation, it follows that student evaluation cannot be made part of official evaluation records for obvious reasons: students are not professional employees, cannot be bound by professional ethics, and cannot be expected to evaluate with professional competency.

Another argument exists to support the case for a code of professional ethics and due process procedure: many problems within a professional group arise not from bad intent or ineptitude but from differences of perception about the nature and the standards of the profession. A code of ethics and a due process procedure may serve as a convenient base from which to deal with many such

difficulties. A rough draft statement of ethics and due process* is attached for consideration.

*This material was not prepared by all of the same people involved in the present study and report concerning evaluation. It is a rough draft of a manuscript still undergoing study and modification.

ADDENDUM TO EVALUATION REPORT

Code of Ethics (Rough Draft)

Preamble

Educators are members of a profession which stresses service and leadership within the community.

Educators, whatever their positions within the profession, subscribe to a code of ethics designed to stimulate exemplary performance.

Educators, sharing the same burden as other professionals, should be self-policing. They are expected to act wisely for the benefit of their students, their colleagues and their community.

Educators agree that resolution of their problems should be made through deliberate exercise of reason, with equity always foremost, and with reference to law and the rules and regulations of the institution wherein the problems arose.

Statement of Ethics

1. The educator's primary interest is the discovery of truth. He is objective and dispassionate in his examination of evidence and respects the opinions of students and colleagues.

2. The educator is obligated to help his students become well-informed and to think for themselves. He stimulates the free pursuit of knowledge within his field of competence and assignment. He rejects the persistent intrusion of material which, in his judgment, has no relation to the subject of his courses.

3. The educator recognizes the sometimes sensitive nature of his relationship with students and colleagues and accepts the principle of confidentiality. Furthermore, he avoids exploitation of students and colleagues for private gain and/or partisan interest.

4. The educator, whatever his position within his institution, has considerable power over others. He exercises that power prudently, knowing his capacity to affect the lives of students and colleagues.

5. The educator abides by the reasonable rules and regulations of his institution and works in an orderly manner to effect changes which he believes desirable. He accepts his share of responsibility to cooperate with colleagues in the community of scholars and in the equitable governance of his institution.

6. The educator subscribes to principles of academic freedom for students and colleagues alike. He recognizes that academic freedom is a sham if denied to anyone and if practiced without responsibility.

7. The educator accepts his responsibility for explaining and interpreting his professional principles and actions when reasonable questions arise, whether from colleagues, students, community or the profession as a whole.

8. The educator recognizes that his actions may easily reflect credit or discredit upon his institution. He conducts himself, therefore, with decorum. He is free to act, write and speak in exercise of his constitutional rights and personal beliefs, but he is careful to distinguish between his activities as a private citizen and as a representative of his institution.

9. The educator respects the principles and practices of due process operant in his institution. He seeks justice for others in the same spirit that he seeks it for himself.