

AN ABSTRACT OF THE THESIS OF

DAVID KENT MAXWELL for the DOCTOR OF EDUCATION  
(Name) (Degree)  
in Education presented on July 23, 1973  
(Major) (Date)

Title: INSTRUCTOR PERFORMANCE APPRAISAL  
IN COMMUNITY COLLEGES

*Redacted for Privacy*

Abstract approved: \_\_\_\_\_  
Dr. Donald Beringson

This study was designed to determine instructor performance appraisal methods being used in community colleges (in Florida, Iowa, New York, Texas, and Washington) and to identify factors responsible for variations in performance appraisal procedures. Questions about instructor performance appraisal which were examined included: who does it, for what purposes, what methods are used, and what is evaluated. Hypotheses which were tested stated that significant differences do not occur (1) among the states, (2) between the responses of administrators and instructors, (3) between the responses from multi-campus and single-campus colleges, and (4) among the various interaction combinations of state, personnel, and campus factors.

## Procedures

Data were obtained with a mailed questionnaire sent to administrators and instructors selected at random from faculties in participating colleges. A total of 160 questionnaires were used, 32 from each state. The data were statistically analyzed through the multi-analysis of variance technique.

## Conclusions

Conclusions were that:

1. The instructor's immediate supervisor is primarily responsible for appraisal, but contributions are provided by students and other administrators.

2. Important purposes for appraisal include: improvement of instruction, to encourage retention of effective personnel, and articulation of goals and objectives which affect supervisor and instructor.

3. Rating scales are the most extensively used method for appraisals.

4. Criteria for measuring instructor performance emphasized: classroom interaction with students; attitude, judgement, initiative, leadership; classroom management; instructional planning; commitment to institutional goals;

interaction with faculty; and improvement since last appraisal.

5. Significant differences occurred among responses from the five states for 25 of the 60 items on the questionnaire.

6. Significant differences between administrator and instructor responses occurred for 19 of the 60 items.

7. Multi-campus and single-campus colleges were essentially not different (the null hypothesis was rejected in four instances).

8. The null hypothesis stating that there are no significant interaction differences for the mean scores of the variables in the study was rejected in 23 instances, but no pattern of rejection was identified.

Instructor Performance Appraisal  
in Community Colleges

by

David Kent Maxwell

A THESIS

submitted to

Oregon State University

in partial fulfillment of  
the requirements for the  
degree of

Doctor of Education

June 1974

APPROVED:

*Redacted for Privacy*

---

Professor of Education \_\_\_\_\_  
in charge of major

*Redacted for Privacy*

---

Dean of School of Education

*Redacted for Privacy*

---

Dean of Graduate School

Date thesis is presented \_\_\_\_\_ July 23, 1973

Typed by Opal Grossnicklaus for David Kent Maxwell

## ACKNOWLEDGEMENTS

This study was made possible by the assistance and cooperation of many people. Special thanks are extended to Dr. Donald Beringson for serving as chairman of the researcher's doctoral committee. Other members of the committee also provided helpful suggestions and guidance. Members of the committee included: Dr. Pat Atteberry, Dr. E. Wayne Courtney, Dr. Jack Rettig, and Dr. Austin Walter.

Community college administrators and instructors were helpful in supplying data for the study. Dr. Charles Carpenter, Dr. Arnold Heuchert, Ms. Belle Kiersky, Dr. Royce Smith, Dr. O. R. Adams, Mr. James Moran, Mr. Glen Nielsen, and Mr. Thor Nielson assisted in refinement of the research instrument.

Dr. Delbert Shirley assisted in the initial planning and organization of the study prior to Dr. Beringson's assignment as committee chairman, and continued to serve in an advisory capacity as the study progressed. Ms. Belle Kiersky edited with persistence and understanding. Library and computer center staff provided timely professional assistance.

Special thanks also go to Kent, Dennis, Deanne, and Faye for their endurance and their clerical assistance. H. F. F. certificates will be awarded.

The cooperation and kindness of everyone who assisted with this study is sincerely appreciated.

## TABLE OF CONTENTS

<u>Chapter</u>	<u>Page</u>
I	INTRODUCTION <span style="float: right;">1</span>
	Purpose of the Study <span style="float: right;">2</span>
	Objectives <span style="float: right;">2</span>
	Need for the Study <span style="float: right;">3</span>
	Hypotheses of the Study <span style="float: right;">8</span>
	Definitions <span style="float: right;">9</span>
	Assumptions <span style="float: right;">11</span>
II	REVIEW OF RELATED LITERATURE <span style="float: right;">13</span>
	Introduction <span style="float: right;">13</span>
	Participants in Instructor Performance <span style="float: right;">14</span>
	Appraisal <span style="float: right;">14</span>
	Purposes for Appraisals <span style="float: right;">22</span>
	Methods of Conducting Performance <span style="float: right;">26</span>
	Appraisals <span style="float: right;">26</span>
	What Has Been Appraised <span style="float: right;">29</span>
	Effectiveness of Appraisals <span style="float: right;">31</span>
	Summary <span style="float: right;">35</span>
III	DESIGN OF THE STUDY <span style="float: right;">37</span>
	Design of the Instrument <span style="float: right;">37</span>
	Selection of the Sample <span style="float: right;">39</span>
	Collection of Data <span style="float: right;">42</span>
	Analysis of Data <span style="float: right;">43</span>
IV	ANALYSIS OF DATA <span style="float: right;">46</span>
	Demographic Data <span style="float: right;">46</span>
	Results of Hypothesis Tests <span style="float: right;">49</span>
	Summary of Means and Hypothesis Tests <span style="float: right;">57</span>
	Data Describing Classroom Observations <span style="float: right;">66</span>
V	CONCLUSIONS, IMPLICATIONS, AND SUGGESTIONS FOR FURTHER STUDY <span style="float: right;">69</span>
	Conclusions <span style="float: right;">69</span>
	Implications <span style="float: right;">75</span>
	Suggestions for Further Study <span style="float: right;">80</span>

	<u>Page</u>
BIBLIOGRAPHY	82
APPENDICES	
APPENDIX A	Questionnaire 87
APPENDIX B	Letter to Presidents 88
APPENDIX C	Sample of Postal Card Sent to Presidents 89
APPENDIX D	Letter to Persons Supplying Directories 90
APPENDIX E	Follow-up Letter to Presidents 91
APPENDIX F	Cover Letter Which Accompanied Questionnaire 92
APPENDIX G	Follow-up Letter to Go With Questionnaire 93
APPENDIX H	Coding of Data Cards 94
APPENDIX I	Comments by Respondents 95
APPENDIX J	Sample Computation of L. S. D. Test 101



## LIST OF TABLES

<u>Table</u>		<u>Page</u>
1.	Responses to Requests for Permission to Include Colleges in This Study	41
2.	Distribution of Questionnaires Used for Statistical Analysis	42
3.	Analysis of Variance	44
4.	Respondents Indicating No Appraisal Process	47
5.	Participants in Formal Evaluation of Instructor Performance	48
6.	Computed F-test Results for Each Factor Tested	50
7.	Means for the States, Personnel, and Campus Factors	53
8.	Results of Least Significant Difference Tests	56

# INSTRUCTOR PERFORMANCE APPRAISAL IN COMMUNITY COLLEGES

## CHAPTER I

### INTRODUCTION

Instructors are appraised by students, parents, other instructors, administrators, and even the public. The format of the system and the type of evidence that is collected or analyzed varies from institution to institution, but evaluations are inevitable.

Instructor appraisal can run the gamut from opinion to precise measurement. Appraisals are made to:

1. Improve teaching.
2. Reward superior performance.
3. Supply information for modifying assignments.
4. Protect either the individual or the organization in legal matters.
5. Provide the basis for planning individual growth and development. (Bolton, 1972)

According to Bolton, educators believe the most important purpose of evaluation is to improve instruction.

As "teaching institutions" community colleges have shared responsibility for providing superior instruction. (Boyer, 1970)

This study has concentrated on determining how community colleges

identify superior instruction and, more specifically, how instructor performance has been appraised.

### Purpose of the Study

The purpose of this study was twofold: (1) to determine the performance appraisal methods being used in community colleges in five states (Florida, Iowa, New York, Texas, and Washington) and (2) to identify factors that account for variations in performance appraisal procedures.

### Objectives

The objectives of this study were to answer the following questions:

1. Who was involved in the appraisal of instructor performance?
2. What purposes did instructors and administrators report as reasons for conducting performance appraisals?
3. What was evaluated and what methods were used in the appraisal of instructor performance?
4. What differences existed in the appraisal procedures used in different states?
5. Did administrators perceive the appraisal process differently than did instructors?

6. Did the appraisal procedures in multi-campus community colleges differ significantly from procedures in single-campus community colleges?

#### Need for the Study

Lloyd Woodburne, a former instructor and administrator at the University of Michigan, studied personnel programs in 46 colleges and universities. (Woodburne, 1950) His study indicated that personnel policies in higher education were non-scientific, ambiguously stated, nebulous, and frequently ineffective. Caplow and McGee, professors at the Universities of Minnesota and Texas, respectively, observed:

It is often impossible for a faculty member to discover his relative position, the opinion which his superiors have of him, the recommendations which have been made concerning his future, or the criteria on which his current performance is being evaluated. (Caplow and McGee, 1958, p. 230)

More recently, institutions such as Oregon State University, Harvard University, the University of Michigan, Columbia University, Cornell University, Duke University, St. Louis University, the University of Rochester, New York University, the University of Pittsburgh, and Northwestern University were criticized because of non-compliance with equal employment opportunity legislation. (Schenck, 1973; Shapely, 1972) Olson (1972) indicated that the same

problems are present in two-year colleges but possibly to a lesser extent. Problems such as this reflected the inadequacy of personnel administration in higher education and indicated a need for a sound system of faculty performance appraisal.

Growth also has led to the need for developing up-to-date methods for evaluating instructor performance in community colleges. Since 1962, the enrollment in Oregon community colleges has increased more than ten-fold. (Langley, 1972) Enrollment in Iowa community colleges since the inception of the statewide area school system in 1966 has more than doubled. (Iowa, 1972) Nationally, enrollments have increased at a rate of approximately 15 percent per year over the last decade. In the midst of such growth, community colleges have been confronted with problems of staff recruitment, selection, and retention. Application of proven personnel management techniques has become increasingly essential.

Blocker, Plumber, and Richardson (1965) suggested that the administration of the community college would be improved if the conventional line-staff organizational structure was modified. This would place more direct emphasis upon educational and personnel functions of the college and assign personnel and administrative responsibilities in a more meaningful way. The specialization they referred to indicated the need for a personnel officer or department at each community college. Bolton (1970) mentioned personnel

directors, students, peers, administrators, and consultants as possible participants in the appraisal process. Little information was available to describe the responsibilities of these personnel in the appraisal process.

Part of the process of reviewing literature on this topic included a DATRIX search (Direct Access to Reference Information, a Zerox Service). DATRIX searches Dissertation Abstracts, and the search conducted for this study revealed that since 1967 a variety of studies related to personnel management in community colleges have been conducted. Six studies concentrated on factors which affect job satisfaction. A dozen studies examined the roles of community college faculties, including roles in the decision-making and policy-formulation processes. The one type of study which was absent from those revealed by the search was a study to examine the procedures used for assessing faculty personnel performance.

Personnel appraisal is a part of MBO (Management By Objectives). Odiorne described management by objectives as a process in which:

The superior and subordinate managers of an organization jointly define its common goals, define each individual's major areas of responsibility in terms of the results expected of him and use these measures as guides for operating the unit and assessing the contribution of its members. (Odiorne, 1965, p. 55-56)

Lahti (1970) indicated that MBO in educational institutions emphasizes management planning and problem solving. He indicated that the process can be used to integrate individual needs for growth and development with organizational goals and objectives in order to make a more effective organization. Applications of MBO at William Rainey Harper Junior College (Illinois) were described, but no data was given to indicate the prevalence of MBO or other methods for appraising instructor performance in similar community colleges.

Wisgoski discussed the importance of appraisal and the lack of information about the methods being used. He stated:

There is almost universal agreement that the single most important ingredient determining the success or failure of any educational institution is the quality of its instructional staff. In higher education this is especially true of the community junior college where emphasis is on teaching rather than research. But how is instruction to be evaluated? While debate on techniques for evaluation have filled educational journals, there is still no agreement on the validity of various methods, and not much information on the degree to which they are utilized. In fact, the evaluation of instruction remains one of the crucial problems facing educational institutions. (Wisgoski, 1970, p. 3)

A state-by-state analysis of teacher evaluation was made by McPhail (1967). He reported that states have attempted to evaluate teachers on the basis of experience and degrees gained. Other evaluative criteria have included National Teacher Examinations (South Carolina and Georgia) and various forms of merit rating systems (New York, Tennessee, North Carolina, Florida, Utah, and Delaware).

It was reported that only the programs in South Carolina, Georgia, Tennessee, and Utah were still in operation. Even in these states the systems of evaluation have generally shifted to local systems of evaluation and away from evaluation by testing and/or statewide standards. Others were discontinued primarily because test results (National Teacher Examination and/or Graduate Record Exam) did not effectively identify teachers who deserved merit raises. McPhail reported variations in statewide evaluations but did not discuss evaluations within individual school systems in the various states. No state-by-state comparisons of instructor performance appraisal in community colleges was discovered, but McPhail's findings suggested that variations exist.

A doctoral dissertation by Burge (1972) was addressed to the development of a system of educator evaluation in the public school system of Eugene, Oregon. School board manuals and educational literature stressed the importance of the relationship between teacher evaluation and the improvement of instruction, but Burge pointed out that school administrators must often supervise and evaluate teachers without the benefit of supporting personnel. Burge explained that it is not uncommon for teachers to question the contribution of administrative evaluations to the improvement of their classroom instruction. Administrators apparently are not cognizant of the attitudes teachers have toward the reliability and validity of administrative evaluations.



The NEA (National Education Association) reported that more than three-fourths of the superintendents and principals responding to a nationwide study by the NEA expressed confidence in their school system's evaluation of teaching, but over half of the teachers did not. (National Education Association, 1964) No evidence was offered to demonstrate that these differences in perception of evaluation also existed between community college administrators and instructors.

The NEA also reported that personnel in small schools, where teacher evaluation follows less formal patterns, were less satisfied with evaluation procedures than personnel in larger systems. No comparison was made between multi-campus and single-campus community colleges, but the possibility of differences may have been suggested since multi-campus colleges tend to be larger than single-campus colleges.

The abundance of information about teacher evaluation has tended to report either isolated examples of the appraisal process or examples which centered on appraisal systems used in universities or in elementary/secondary schools. This study will focus on the appraisal process at the community college level.

### Hypotheses of the Study

The following series of null hypotheses were tested for sources of variation in personnel performance appraisal procedures:

H<sub>1</sub> There is no significant difference among the mean scores of responses from the five states included in the study.

H<sub>2</sub> There is no significant difference between the mean scores of responses from administrators and instructors.

H<sub>3</sub> There is no significant difference between the mean scores of responses from multi-campus and single-campus institutions.

H<sub>4</sub> There are no significant interaction differences for the various mean score interactions of responses representing the state, personnel, or campus variables included in the study.

### Definitions

The following terms were defined to clarify their meaning as used in this study. Other terms in the text of the study were considered to be self-explanatory.

Administrator: A community college faculty member whose major assignment consists of superintending the resources of the college, including personnel. For this study administrators directly involved in the instructor appraisal process were selected (deans of instruction, division chairmen, and others with comparable responsibilities).

Community College: A public post-secondary school established for the purposes of providing courses of study limited to not more than

two years full-time attendance and designed to meet the needs of a geographical area by providing educational services, including but not limited to vocational-technical programs and lower division collegiate programs. The comprehensive nature of the institution was its identifying characteristic, and other names such as junior college, area community college, community junior college, or other combinations were used in some communities.

Critical incidents: A notebook record of incidents that indicate effective and ineffective actions by the instructor.

Field review: The instructor's supervisor is interviewed by a personnel officer or administrator and the interviewer prepares a report. The instructor is not interviewed.

Group appraisal: The instructor's supervisor and other administrators or personnel officers who are familiar with the instructor's work have a conference in which they collectively appraise the instructor's performance.

Instructor: A community college faculty member whose major assignment is teaching.

Multi-campus college: Community colleges which have an administrative office serving more than one campus. Extension centers, remote classrooms, and other off-campus facilities constituted a second campus only when both transfer and vocational-technical programs were available. American Junior Colleges (Gleazer, 1971)

and Opportunities in Iowa's Area Schools (Iowa, 1972) were used as authorities to identify multi-campus colleges.

Performance appraisal: Systematic evaluation of an instructor with respect to his performance on the job and his potential for development. Other terms commonly used in lieu of performance appraisal include: personnel appraisal, employee appraisal, personnel review, progress report, service rating, performance evaluation, and fitness report. (Beach, 1970)

Single-campus college: American Junior Colleges (Gleazer, 1971) and Opportunities in Iowa's Area Schools (Iowa, 1972) were used to identify single-campus colleges. Such colleges generally had administrators operating out of a single location even though instruction and supervision of instruction occurred in off-campus centers.

### Assumptions

Assumptions related to the procedures used in this study included:

1. Generalization of findings was made only within the limits of the populations surveyed. Only five states were sampled, and the population included only administrators and instructors in those community colleges where presidents consented to have a sample of their faculty surveyed. Presidents, deans of students, registrars, and directors of counseling were

not included in the population of administrators from which the sample was drawn. Their responsibilities generally included very little involvement in instructor performance appraisal. Deans of instruction, division chairmen, and other administrators were utilized for the purposes of this study.

2. It was assumed that the instrument used for the survey was designed to elicit responses which accurately reflected instructor performance appraisal procedures in the community colleges surveyed. That assumption was supported by a field test which indicated that the instrument could secure accurate data. The underlying assumption was that the survey would produce accurate information about the population sampled.

## CHAPTER II

## REVIEW OF RELATED LITERATURE

Introduction

Instructor performance appraisal was described by the Florida Inter-institutional Research Council (Schafer, 1971) as a process involving four basic steps:

1. Definition of the objectives of the institution.
2. Identification of the role of faculty in accomplishing the objectives (job analyses and job descriptions).
3. Definition of acceptable standards of performance.
4. Collection and interpretation of evidence in regard to the standard.

Appraisal procedures involve many facets of the operation of the institution including institutional philosophy. Complications include the ambivalence and frustration which has permeated accumulating research aimed at answering questions such as:

1. What are the characteristics of an effective teacher?
2. What are effective teacher-learner relationships?
3. What are effective instructional techniques?
4. What are the criteria of effective teaching?
5. How can the factors which affect the outcomes of teaching be defined and controlled? (Harsh, 1970)

Conclusive answers to these questions have not been found, but the need to make decisions about instructor performance has continued.

This review focuses on selected samples of the literature and research on instructor performance appraisal. Specific questions for which answers were sought include: who is involved in the process and to what extent, why appraisals are performed, what methods are used, what qualities are evaluated, and are appraisals effective?

### Participants in Instructor Performance Appraisal

#### Colleagues

Colleague evaluations were mentioned by various writers, but research on the subject was found to be rather limited. This was partially explained by the limited opportunities instructors have for inter-class visitations. Howsam commented:

It would appear that there is little to gain through the use of peer ratings unless either teachers are given opportunities to observe their peers at work or the assessment areas are limited to those for which there is opportunity to observe (work on committees, participation in staff affairs, assistance to other teachers). (Howsam, 1963, p. 16)

Another possible explanation was the reluctance of instructors to evaluate or rate one another.

Fattu (1963) observed that studies have shown substantial agreement in the ratings of teachers when the rating was done by

supervisors and by fellow teachers. The suggestion was that halo effects influence peer ratings just as they do administrative ratings. (The term "halo effects" refers to a tendency to give high ratings to individuals with personal qualities the evaluator admires.)

Morsh, Burgess, and Smith (1956), after a study involving military instructors, concluded that instructors generally base their opinions of one another on such evidence as knowledge of subject matter rather than on the actual effectiveness of their teaching. Along similar lines, McCall (1952) suggested that colleagues who get better results in terms of pupil gain may be judged by fellow teachers to be poorer teachers.

### Student Evaluations

Students were found to be promising but little used sources of information. Research reviewed by Howsam (1963) indicated that:

1. Students tend to agree with one another in their ratings of instructors.
2. Instructors rated highest by students tend to enable students to learn more.
3. Student ratings often differ from those made by supervisors and other teachers.



Howsam summarized student evaluations by saying:

With remarkable consistency, the findings have shown that pupils are able to make more valid and reliable ratings of teachers than any other group, including administrators, supervisors, and experts. Teachers in these studies have found pupil ratings to be both fair and accurate. Despite the favorable evidence there is widespread resistance to the rise of pupil ratings, probably arising out of the respective roles of student and teacher in our culture. (Howsam, 1963, p. 16)

Research summarized in "Junior College Research Review" (Boyer, 1970) noted that student ratings of instructors are not substantially related to the student's age, sex, grade point, or grade(s) previously received from the instructor being rated.

The research findings of Remmers (1963) supported Howsam's conclusions and also provided the following generalizations:

1. Reliability ratings of teachers by students is directly related to the number of raters.
2. Little if any relationship exists between the difficulty of a course and the student's rating of an instructor.
3. Alumni, 10 years after graduation, agree substantially with on-campus students in their average rating of the same instructors.
4. The instructor's popularity in extra-class activities is not appreciably related to student ratings of that instructor.
5. Students are more favorable than instructors to student appraisal of instructors, but more instructors than students

have noticed improvement in their teaching as a result of student ratings.

Researchers pointed out that there are factors not directly related to the instructor's performance, such as class size, which may affect student ratings. (Cohen, 1969) Even so, the available evidence suggested that student evaluations are a relatively untapped and quite reliable source.

### Administrative Evaluations

Observations and evaluations by administrators constituted the most widely used measure of teacher competence. (Hain and Smith, 1968) It was reported that use of administrative evaluations continued even though various research findings had reported that such ratings were apt to be contaminated with halo effects (Fattu, 1963; Morsh, Burgess, Smith, 1956) and for the most part do not produce very high correlations with student gain. (Morsh, Burgess, and Smith, 1956)

Rose (1963) studied five school systems that were seriously attempting to develop better programs of teacher evaluation. He reported that principals face the following practical problems as teacher evaluators:

1. Taking the time necessary to do an adequate job.
2. Developing skills in communication and consistent follow-through with teachers.

3. Varying value systems and interpretive criteria for the assessment of data about teaching.

These findings were similar to observations made by Barr:

1. There is plenty of evidence to indicate that different practitioners observing the same teacher teach, or studying data about her, may arrive at very different evaluations of her; this observation is equally true of the evaluation expert; starting with different assumptions, employing different approaches, and using different data-gathering devices they, too may arrive at very different evaluations, and

2. Most evaluators attempt to make judgments about small differences in effectiveness that do not seem to be possible at the present time, considering current know-how and data-gathering devices. Possibly for the time being it might be best to attempt to set up only broad categories of teacher effectiveness, such as adequate, superior, and inadequate, and to do this with reference to pretty carefully defined situations. (Barr, 1961, p. 150-151)

Bolton posed the question: can administrators also supervise?

When an administrator must help his teachers improve and at the same time be accountable to the school system for results, he may have the tendency to emphasize his responsibility for accountability, thus reducing his effectiveness and trust relationship with his staff.

Bolton analyzed this apparent dichotomy as follows:

The issue is not whether those who supervise and provide assistance should evaluate teachers; rather, it is how information obtained through the evaluation process should be used. (Bolton, 1970, p. 42)

Rose (1963) took a positive approach to the issue and observed that: (1) an adequate evaluation program should ameliorate these

aspects of organizational life, and (2) principals (supervisors) in systems of any size should not be required to make personnel decisions (salary, tenure, dismissal, etc.) directly and without the involvement of the superintendent and his staff.

Writers and practitioners saw the principal's (supervisor's) role in teacher evaluation as being filled with hazards and problems, and few solutions have been perfected. The question of whether an administrator can effectively evaluate and simultaneously assist teachers with their instructional problems has not been answered.

#### Evaluations by Consultants and Others

Bolton (1972) indicated that subject matter specialists, consultants, and parents should be considered when developing plans for instructor performance appraisals. State departments and advisory committees may also assist with the analysis of an instructor's work.

#### Self-Evaluations

Writers and researchers in the field of education tended to disagree about the value and place of self-evaluations. The National Education Association (1964) reported the results of a study in which 19.2 percent of the teachers reported being required to complete written self-evaluations as a part of their school system's evaluation process. The practice was reported more frequently in large

systems than in small systems and in elementary schools more than in secondary schools.

Howsam, summarizing his views of research about self-ratings, stated: "There is little, however, in either personality theory or teacher competence research to indicate that valid or reliable assessments come out of such efforts." (Howsam, 1963, p. 16) He did acknowledge that a limited amount of research by the California Council on Teacher Education indicated otherwise. Howsam suggested that the differences in the research findings may be due to the climate within which the ratings took place. If the self-rating is part of a research project, the teacher knows that his identity will not be known and is likely to be more objective. If self-rating is part of the school system's formal evaluation process, the teacher is inclined to over-rate himself and not to report incriminating data.

Musella recognized weaknesses in self-rating systems and suggested that evaluation plans focusing on self-improvement should not become a part of the institutional requirements for assessing teacher competence. He suggested there should be an external (institutional) and an internal (self) system of accountability.

1. One that satisfies the institutional needs for perpetuation and purpose achievement; and

2. One that satisfies the needs of improvement of process for both the individual and the institution.  
(Musella, 1970, p. 15)

Other disadvantages of self-ratings were that such ratings did not relate to an outside criterion and that teachers were not able to analyze specific aspects of their behavior because they were without a conceptual framework or systematic observation system. (Bolton, 1970) Bolton recommended that the following preparatory activities be completed before implementing a self-evaluating approach for improving instruction:

1. Teachers should be provided with a framework (an observational system) for analyzing and interpreting their own behavior.

2. Teachers should be provided with the technical competence needed for operating the various new media used for recording their behavior. (Bolton, 1970, p. 167)

Performance goals as discussed by McGregor (1960) and self-evaluation systems as proposed by Bolton were found to be similar.

Both stressed the following:

1. The rater and ratee develop criteria against which certain behaviors can be assessed.

2. The rater and ratee focus on actual events and specific behaviors.

3. The situation remains relatively free from external threat.

4. The rater and ratee have a common frame of reference for judging effectiveness minimizing individualized value judgments.

Perhaps the concept was summarized by the TEPS (Teacher Evaluation and Professional Standards) commission in a statement on teacher evaluation. They stated that self-evaluation is basic. (National Education Association, 1964)

### Purposes for Appraisals

Bolton systematically searched the literature in the behavioral sciences and surveyed school districts throughout the United States on the topic of personnel evaluation. He stated:

One of the first steps in establishing or revising a program of teacher evaluation is the determination of purposes of the program. If the program is to be successful, it is necessary that these purposes be identified, discussed, and agreed upon by all who are involved in the process. All potential purposes should be discussed openly and thoroughly; certain purposes should not be ignored just because there are critical issues involved in them or because their accomplishment may precipitate conflict. (Bolton, 1970, 21-C)

The purposes which Bolton spoke of were:

1. To improve teaching, including out of classroom activities. . . (This purpose is not limited to teacher behavior but implies any actions taken to improve teaching systems, the teaching environment, or teacher behavior.)
2. To reward superior performance.
3. To supply information for modification of assignments (including placement in another position, reduction of load, promotion to a leadership position, or termination . . .)

4. To protect individuals or the school system in legal matters (protection of teachers against a capricious new administrator and the protection of the school district and children against a harmful teacher).

5. To validate the selection process.

6. To provide a basis for career planning and individual growth and development of the teacher. (Bolton, 1970, 21-C)

Howsam (1963) explored the problem of identifying and evaluating effective teaching and suggested that the following additional goals be included among the purposes for evaluations:

1. To determine the effectiveness of the instructional program. Teaching is only a means to an end, but it may be possible and necessary to infer the achievement of ends through effective means.

2. To determine the effectiveness of personnel policies and procedures and to obtain and retain quality teachers.

3. To provide the basis for supervisory and inservice development programs and activities. Programs which are aimed at on-the-job improvement should grow out of the particular needs of individuals and groups and be based on objective evidence.

4. To provide one form of evidence that can be used as the basis for administrative decisions on personnel and programs.



5. To facilitate accounting for responsibility. Teachers are given authority and responsibility and must be accountable. Evaluations are used by administrators to report their stewardship.

6. To motivate teachers to strive for a high level of performance. Some believe that the mere presence of an evaluative procedure provides motivation to a greater effort.

7. To provide the basis for rewards or sanctions. Merit pay is but one example of this.

8. To assist the teacher in achieving success. This may be for institutional purposes or directed towards the teacher.

Redfern, a personnel director with experience in implementing an evaluation plan in a large city school system, suggested that the purposes of evaluation vary from school system to school system.

Purposes include:

Assessment of the status and quality of teaching performances; identification of these aspects of performance which are below standard and need improvement; and stimulation of the growth and development of the individual. (Redfern, 1963, p. 25)

Additional incentive for appraising the performance of instructors was found to exist in the form of state laws. In Iowa provisions must be made to "inspire improvement of instruction, development of general policy, . . ." (Iowa, 1971, p. 17) In practice Iowa community colleges are expected to be able to demonstrate that personnel

performance appraisals are made. More than a decade ago, the Florida state legislature made it mandatory for state and county superintendents to maintain a personnel file for each certificated teacher. An evaluation of the teacher's services as a teacher was to be included in that file. (National Education Association, 1964)

The NEA (National Education Association) pointed out the professional responsibility to maintain standards of quality for members of the profession. NEA resolutions every year since the late 1950's have stated that it is a major responsibility of the teaching profession to evaluate the quality of its services. The NEA also pointed out that citizens were demanding merit rating in which superior teaching could be rewarded and inferior teaching penalized.

Most authorities tended to agree on the purposes for conducting instructor performance appraisals and that the results of evaluations must be translated into effective supervision programs before a significant improvement in performance is likely to result. Burge observed:

Unfortunately, supervisory activities which should be used to improve teaching have frequently been regarded with suspicion, and as a result, supervision--specifically classroom visitation and observation--has fallen into disrepute. (Burge, 1972, p. 24)

### Methods of Conducting Performance Appraisals

Dale S. Beach, a Professor in the School of Management at Rensselaer Polytechnic Institute (Troy, New York) said that over the years a considerable number of appraisal systems have been developed and an effort has continued to make them more objective, more valid, and less dependent upon unsupported whims of the raters. He identified the following major types of rating systems:

1. Rating scales.
2. Employee comparison.
  - a. Ranking.
  - b. Forced distribution.
3. Check list.
  - a. Weighted check list.
  - b. Forced choice.
4. Critical incident.
5. Field review.
6. Free-form essay.
7. Group appraisal.
8. Appraisal by results. (Beach, 1970, p. 316)

Educators have used these methods in various combinations and with various modifications. In Teacher Performance Evaluation Procedures for instance, the Oregon Board of Education recommends evaluation forms which use rating scales. Other forms include:

checklists, essay-type responses, and work sheets which apply the MBO (Management By Objectives) concept. (Oregon Board of Education, 1971)

In the field of personnel management Douglas McGregor coined the Theory X and Theory Y labels when he contrasted assumptions about human behavior which are commonly found in the field of business management. In The Human Side of Enterprise, he stated that under Theory X management makes these assumptions about behavior:

1. The average human being has an inherent dislike of work and will avoid it if he can. . . .
2. Most people must be coerced, controlled, directed, threatened with punishment to get them to put forth adequate effort toward the achievement of organizational objectives. . . .
3. The average human being prefers to be directed, wishes to avoid responsibility, has relatively little ambition, wants security above all. (McGregor, 1960, p. 33-34)

McGregor stated that the assumptions of Theory X are giving way to the newer concepts of Theory Y under which it is assumed that:

1. The expenditure of physical and mental effort in work is as natural as play or rest. . . .
2. External control and the threat of punishment are not the only means for bringing about effort toward organizational objectives. Man will exercise self-direction and self-control in the service of objectives to which he is committed.
3. Commitment to objectives is a function of the rewards associated with their achievement. . . .
4. The average human being learns, under proper conditions, not only to accept but to seek responsibility. . . .

5. The capacity to exercise a relatively high degree of imagination, ingenuity, and creativity in the solution of organizational problems is widely, not narrowly, distributed in the population.

6. Under the conditions of modern industrial life, the intellectual potentialities of the average human being are only partially utilized. (McGregor, 1960, p. 47-48)

McGregor proposed that each subordinate establish short-term performance goals for himself. The supervisor's role was seen as helping the subordinate relate his self-appraisal, his job targets or performance goals, and his plan for the ensuing period to the realities of the organization. In this process emphasis shifts from appraisal to analysis. It is assumed that the individual knows--or can learn--more than anyone else about his own capabilities, needs, strengths, weaknesses, and goals. The purpose of the plan is to establish realistic targets and to seek the most effective ways of reaching them. Thus appraisal becomes a means to a constructive end.

McGregor's approach, applied to educator evaluation, would place the major responsibility on each educator for establishing performance goals and appraising progress toward them, thus stimulating self-development and self-evaluation. It seemed probable that such an approach would require more supervisory skill and time, but McGregor contended that the greater motivation and the more effective development of subordinates would justify the added costs.

### What Has Been Appraised

Caplow (1958) reported that at the university level research, publishing, and teaching were among the primary items to be appraised. He reported that out of 371 responses 122 respondents defined "productivity" as research, only 14 referred directly to the teaching of students, and 235 of the responses were so worded that the criteria used to describe "productivity" were obscure.

Bolton (1972) pointed out that schools are established to facilitate pupil learning; therefore, the ultimate criterion for teacher success is the amount of learning that occurs in the students for whom he is responsible. Measurement of student growth was reported to have the advantage of being a direct measure of the outcomes that are desired. Disadvantages were reported to be centered around difficulties in (1) determining how much of the growth can be attributed to a given teacher, (2) determining the teacher's effect on both short-term and long-term growth, and (3) accounting for the multitude of uncontrollable variables. Bolton reported that, historically, student accomplishment has been avoided as a means for evaluating teachers, but recent emphasis on accountability has caused renewed interest in establishing specific student accomplishment goals and attempting to reach them.

Conversely, a recent decision by a federal judge in Iowa established restrictions on testing of students as a means of evaluating

teachers. The court said, "A teacher's professional competence cannot be determined solely on the basis of her student's achievement." (A Court Decides, 1973) The decision did not rule out the possibility that testing might be included as an element of a more comprehensive system of appraisal.

Flanders (1964) developed a numerical system for analyzing various types of classroom interaction. The system is designed for use in situations in which the teacher and students are actively discussing school work. Flanders demonstrated that certain types of teaching effectiveness can be indicated by categorizing verbal interaction. Other systems of interaction analysis examine response patterns of students, verbal patterns used by the teacher, evaluation of achievement of stated objectives for a specific teaching session, and arrangement of the classroom as a means of eliciting active interaction from all students. (Goldhammer, 1969) These evaluations all require the presence of evaluators in the classroom.

Bolin and Muir (1966), in a study reviewing merit rating literature, observed that important factors involved in faculty rating (in descending order) are: (1) classroom teaching, (2) personal attributes, (3) professional growth activities, (4) student-faculty relations, (5) community service, and (6) research and publication. The importance of these factors was supported by Schafer (1971). Discussing evaluation of faculty in Florida's community junior colleges, he

indicated that the roles of community college instructors include: teaching and advisement of students, committee work, public service, and service to the profession.

### Effectiveness of Appraisals

Appraisal procedures and the effectiveness of those procedures were examined in a nationwide survey of public schools. (National Education Association, 1964) The survey sought answers to the following questions:

1. How prevalent are formal, written plans for the evaluation of classroom teachers?
2. Have criteria for teacher competency been established by school systems?
3. Do evaluation practices vary for probationary and continuing teachers, elementary and secondary teachers, large and small districts?
4. Who is responsible for evaluating teachers?
5. What methods are used to evaluate teachers?
6. What use is made of the evaluation?
7. Do teachers have a voice in the evaluation programs?
8. What opinions are held in regard to the evaluation program by the various persons involved? (National Education Association, 1964, p. 6)

Questionnaires were sent to superintendents, principals, and teachers. Findings of the study indicated that:



1. Large school systems tended to evaluate teachers according to formal rules and regulations while small schools tended to have informal evaluations with no particular rules spelled out.

2. Some schools require a written evaluation which goes to the teacher and to the central office.

3. Formal evaluations are more likely to apply to probationary teachers than to teachers with tenure or on continuing contract.

4. Superintendents had more confidence in their school system's evaluation program than did principals.

5. Teachers had much less confidence in the evaluation program than did superintendents or principals.

6. Personnel in small school systems were less satisfied with teacher evaluations than school personnel in larger systems.

7. Teachers and principals in systems where teachers received a written evaluation the preceding year were "less distrustful" of evaluations than others included in the study.

8. All three groups indicated that evaluation stimulated instruction, led to the development of good rapport between teachers and administrators, led to better administrative planning, and led to identification and release of incompetent teachers.

9. Major criticisms by teachers were that (1) evaluations were not accurate and (2) the administrative staff was too busy to do an effective job of evaluation.

Conclusions drawn from the survey were that effective evaluation programs can be carried on in public schools. To be effective, evaluations must be a high priority among administrative staff. Time must be allowed to prepare people as evaluators, and continuing programs are needed for explaining and interpreting evaluations, especially as new teachers and administrators join the staff.

Along similar lines Bolton recommended periodical assessment of the teacher evaluation process. He stated that an analysis should include examining the realism of the goals of the process and should seek answers to the following questions:

1. Is the instruction improving?
2. Are teachers receiving assistance?
3. Are students learning?
4. Are teachers with problems improving?
5. Are consistently ineffective teachers being released?
6. Do teachers understand what is expected of them?
7. Is adequate information being provided to improve the selection process?
8. Is the board of education provided adequate information for making personnel and policy decisions? (Bolton, 1972, p. 1-2 of 21-1)

Bolton suggested that teachers, administrators, students, and parents may be able to help answer these questions. Exit interviews were recommended for acquiring information regarding:

1. Why individuals leave the organization.
2. The individual's perception of problems in supervision and evaluation.
3. The individual's perception of problems in selection and placement procedures.

Bolton suggested that public school systems treat the training of evaluators more casually than industry does, in spite of evidence that training is likely to increase validity and reliability of decisions and improve discrimination of measurement. Procedures recommended for improving evaluator performance include:

1. Elective inservice courses.
2. University courses.
3. Group meeting devoted to evaluation.
4. General explanations given at regular administrative meetings.
5. Workshops or clinics lasting from 1 to 3 days (including assistance from outside consultants, practice, discussions, use of multimedia presentations).
6. Written documents or manuals.
7. Individual consultation. (Bolton, 1972, p. 4 of 21-I)

### Summary

The literature reviewed in this chapter presented a necessarily abridged summary of the vast amount of writing which has been done on the subject of evaluating the performance of instructional personnel.

The literature has indicated that:

1. Participants in the appraisal process have included: colleagues, students, administrators, instructors (self-evaluations), and consultants.
2. Purposes for appraisal include: clinical or developmental purposes (analysis, planning, improvement of instruction, and motivation) and managerial purposes (for decisions concerning tenure, promotion, merit awards, and legal protection).
3. Methods used include: rating scales, ranking and forced distributions, checklists, critical incidents reporting, field reviews, free-form essay, group appraisal, and management by objectives.
4. Types of performance appraised include: teaching, [amount of] learning by students, classroom interaction, personal attributes, professional growth, student-faculty relations, community service, research, and publishing.

Reasonably effective instructor performance appraisal systems have been developed in some schools, but the time and effort expended for implementation of these systems has been extensive. The fact remains that personnel decisions are made, and because of the personal, legal, and moral implications of those decisions, educators have an obligation to continue to refine and perfect personnel appraisal systems.

## CHAPTER III

### DESIGN OF THE STUDY

This research was an empirical investigation of the relationships between states and between multi-campus and single-campus colleges as perceived by instructors and administrators. The study was designed to determine performance appraisal procedures being used in community colleges in Florida, Iowa, New York, Texas, and Washington. This study also identified factors that accounted for variations in appraisal methods.

#### Design of the Instrument

The survey instrument used for this research was composed primarily of questions adapted from questionnaires used by (1) the National Education Association (1964) in their nationwide study of teacher evaluation and (2) by Burge (1972) in his study of teacher evaluation in the schools of Eugene, Oregon. Adaptations were made to (1) include terminology which was more consistent with that used in community colleges, and (2) to permit responses to the questions to be recorded on a five-point interval scale.

A jury of experts reviewed the instrument to recommend modifications, additions, and deletions they felt necessary. These changes

were made as recommended. The jury was composed of the following persons:

1. Dr. Charles Carpenter, Coordinator of Community College Education at Oregon State University.
2. Dr. Arnold Heuchert, Director of the Career Education Personnel Development Center in Portland, Oregon.
3. Dr. Royce Smith, Assistant Professor of Business Administration at Oregon State University.
4. Ms. Belle Kiersky, Business Instructor at Linn-Benton Community College.

The instrument was field-tested in a medium-sized community college and in a small community college in Oregon. Persons assisting with the field test included:

1. Mr. James Moran, Instructor at Linn-Benton Community College.
2. Dr. O. R. Adams, Dean of Instruction at Linn-Benton Community College.
3. Mr. Thor Nielsen, Instructor at Rogue Community College.
4. Mr. Glen Nielsen, Dean of Instruction at Rogue Community College.

After each of these persons had responded to the items on the questionnaire they were interviewed to determine if the questions were

clearly stated and if the responses given correctly reflected the type of responses sought. A revised copy of the survey instrument reflecting refinements which resulted from the field testing process was prepared for mailing. A sample of the questionnaire is shown in Appendix A.

### Selection of the Sample

The basic arrangement for the study was a  $2 \times 2 \times 5$  factorial design with two levels of personnel, two levels of campus types, and five levels of states. The design represented a fixed model.

The states of Florida, Iowa, New York, Texas, and Washington were selected for the survey because they each had at least three single-campus and three or more multi-campus community colleges, and they represented a geographical cross-section of the country. Permission to include personnel in each of the colleges in these states was obtained from the respective college presidents (superintendents or chancellors) without intervention of a central agency.

Initially, a letter was sent to the presidents of the selected institutions. They were asked to return a self-addressed postal card on which they were to check a box indicating their (1) willingness to participate in the study, or (2) unwillingness to participate. Those who agreed to participate were asked to give the name of a person who could provide a faculty directory which was to be used for



selection of a random sample of instructors and administrators. (See Appendix B and Appendix C.) Some presidents sent faculty directories when they returned the postal card agreeing to participate. The letter in Appendix D was sent to elicit faculty directories from persons identified by the presidents. The request was addressed to registrars when the name of a person to contact was not given.

A follow-up letter was sent to presidents from whom no postal card had been received by January 19, 1973. A sample of that follow-up letter is shown in Appendix E.

Requests for directories were made as affirmative responses from the presidents returned. Two colleges neglected to send directories as requested; therefore, were not included in the study.

Table 1, page 41, shows the number of colleges asked to participate in the study, as well as the number actually participating. A total of 149 colleges were asked to participate, and 52 percent of those (a total of 78) actually participated. The greatest percent of participation was in Iowa where 100 percent of the colleges participated. The least percent of participation was in Texas where only 35 percent of their 34 single-campus colleges participated.

Administrators and instructors were selected at random from faculties in participating colleges in each of the five states on February 23, 1973. A total of 48 persons were selected in each state: 12 multi-campus administrators, 12 multi-campus instructors,

Table 1. Responses to Requests for Permission to Include Colleges in This Study

State and Campus Type	Number Invited to Participate	Did Not Respond	Declined Participation	Participated	Percent Participation
Florida					
Multi-campus	6	1	1	4	67%
Single-campus	22	0	13	9	41%
Iowa					
Multi-campus	8	0	0	8	100%
Single-campus	3	0	0	3	100%
New York					
Multi-campus	9	3	2	4	44%
Single-campus	30	7	11	12	40%
Texas					
Multi-campus	10	1	3	6	60%
Single-campus	34	10	12	12	35%
Washington					
Multi-campus	9	1	2	6	67%
Single-campus	18	3	1	14	78%
Total	149	26	45	78	52%

12 single-campus administrators, and 12 single campus instructors. Four persons in each of the four groups in each of the states served as alternates. A total of 160 of the 240 questionnaires mailed were used when the data were analyzed. The matrix shown in Table 2 (below) illustrates the equal-cell arrangement which was used. Briefly stated, eight questionnaires were selected for analysis from each of the four groups in each of the states.

Table 2. Distribution of Questionnaires Used for Statistical Analysis

State	<u>Multi-campus</u>		<u>Single-campus</u>		$\Sigma$
	Admin.	Instr.	Admin.	Instr.	
Florida	n=8	n=8	n=8	n=8	32
Iowa	n=8	n=8	n=8	n=8	32
New York	n=8	n=8	n=8	n=8	32
Texas	n=8	n=8	n=8	n=8	32
Washington	n=8	n=8	n=8	n=8	32
$\Sigma$	40	40	40	40	$\Sigma\Sigma = 160$

### Collection of Data

Data were collected by mailing a self-addressed, stamped questionnaire to each of the persons selected in the random sampling. The letter in Appendix F accompanied the questionnaire. The follow-up letter shown in Appendix G was mailed three weeks later to those who had not responded.

Data on completed questionnaires were coded and transferred to data processing cards for computer analysis. The coding system is described in Appendix H.

A total of 192 questionnaires, 80 percent of the 240 which were mailed, were returned to the investigator. Of those, 164 were usable. Three were incomplete, 17 indicated no appraisal process was evident, five were from persons selected as administrators who indicated that they were instructors, one was from a person selected as an instructor who indicated that he was an administrator, one was from an administrator who worked only with part-time instructors, and one was returned, undelivered.

#### Analysis of Data

The coded data were electronically processed by the CDC 3300 computer at the Oregon State University Computer Center. The F statistic was used to test each of the hypotheses with the significance level at  $\alpha = .05$ . Table 3, page 44, reflects the significance testing schema. The null hypothesis was considered rejected for each item in which the computed F value exceeded the tabular value listed in the "critical F value" column of the table.

When the hypotheses were tested and retained as a result, no further analysis was made. Where hypotheses were rejected, multiple comparisons analyses were completed for the data. For the states

Table 3. Analysis of Variance

Sources of Variation	df	SS	MS	F	Critical F Value $\alpha = .05$
State	4	A	A/4	$MS_S / MS_{error}$	$F_{4, 140} > 2.43$
Personnel	1	B	B/1	$MS_P / MS_{error}$	$F_{1, 140} > 3.91$
Campus	1	C	C/1	$MS_C / MS_{error}$	$F_{1, 140} > 3.91$
S * x P **	4	D	D/4	$MS_{SxP} / MS_{error}$	$F_{4, 140} > 2.43$
S x C ***	4	E	E/4	$MS_{SxC} / MS_{error}$	$F_{4, 140} > 2.43$
P x C	1	F	F/1	$MS_{PxC} / MS_{error}$	$F_{1, 140} > 3.91$
S x P x C	4	G	G/4	$MS_{SxPxC} / MS_{error}$	$F_{4, 140} > 2.43$
Error	140	H	H/140		
Total	159				

\*Refers to the states factor.

\*\*Refers to the personnel factor.

\*\*\*Refers to the campus factor.

factor, items were statistically analyzed using the Least Significant Difference (L. S. D. ) Test to determine where differences occurred among adjacent ranked mean scores. The  $\alpha$ -level for the L. S. D. tests was set at .05.

Demographic data collected in part A of the questionnaire were used for control purposes and were not statistically analyzed. For Part H of the questionnaire, percentages were calculated to provide an indication of the extent to which classroom observations of the instructor's performance were made and how those visits were conducted.

## CHAPTER IV

### ANALYSIS OF DATA

The findings of this study are presented in four sections which include: (1) demographic data, (2) results of the hypothesis tests, (3) a summary of means and hypothesis tests, and (4) a review of the data describing classroom visitations. Comments by respondents are included in Appendix I.

#### Demographic Data

Data for this study were obtained from questionnaires completed by community college administrators and instructors. Part A of the questionnaire provided demographic data. (See Appendix A.)

The first question in this section was included to assure that administrators and instructors were correctly categorized. Five questionnaires were omitted because they were completed by instructors who were selected in the administrative group. One was omitted because it was completed by an administrator who was selected as an instructor. These errors in selection amounted to 3.2 percent of the 192 persons responding to the questionnaire.

The second question allowed persons completing the questionnaire to indicate the extent of their understanding of the instructor performance appraisal process. Fifty-eight percent of the persons

responding to the questionnaire indicated that appraisal procedures were clearly defined, 33 percent indicated that the process was informal (not clearly defined), and nine percent indicated that no appraisal process was evident. Those responding in the latter category returned the uncompleted questionnaire. Table 4, below, illustrates the origin of those responses.

Table 4. Respondents Indicating No Appraisal Process

State	<u>Multi-campus</u>		<u>Single-campus</u>	
	Admin.	Instr.	Admin.	Instr.
Florida	1	0	0	1
Iowa	1	1	2	3
New York	0	0	0	0
Texas	1	1	1	0
Washington	1	2	1	1
Total	4	4	4	5

Administrators returned 47 percent and instructors returned 53 percent of the responses indicating that no appraisal process was evident. Comparatively, 39 percent of those indicating that the appraisal process was not clearly defined (the second choice in question 2) were administrators and 61 percent were instructors.

The third question was included to determine if all instructors in community colleges were appraised by the same system, and if not, what grouping of instructors for appraisal purposes was



prevalent. (The information in the remainder of this report refers to the 160 questionnaires which were statistically analyzed and omits those responses in the 32 returned questionnaires which were either from alternates or not usable.) A total of 86 percent indicated that all instructors are appraised by essentially the same system, four percent checked vocational-technical, three percent checked lower division, four percent indicated a particular division, and three percent checked "others." Included in the "others" category were those seeking promotion, reassignment, or tenure.

Question four was included to provide an indication of who performs formal appraisals in community colleges. Table 5, below, shows the responses to this question.

Table 5. Participants in Formal Evaluation of Instructor Performance.

Participant	Number of Responses				Overall Total	Overall % of 160
	Multi-campus Admin.	Multi-campus Instr.	Single-campus Admin.	Single-campus Instr.		
Colleagues	12	15	20	15	62	39%
Students	19	17	27	28	91	57%
Personnel Department	1	2	0	1	4	2%
Immediate Supervisor	35	33	28	33	129	81%
Administrators	17	15	13	18	63	39%
Self	16	6	15	13	50	31%
Outsiders	0	0	1	3	4	2%
Others	0	0	3	0	3	2%

The table indicates that the most frequently reported formal input comes from the instructor's immediate supervisor. More than half of the respondents reported that students provide formal inputs.

### Results of Hypothesis Tests

The questionnaire used in this study contained 60 items related to various facets of instructor performance appraisal. Respondents were asked to respond to each item by marking a five-point scale indicating the importance of a characteristic or extent of agreement with a statement. (See Appendix A.) Responses ranged from very slight to very great, with point values of 1, 2, 3, 4, or 5 assigned to each item.

The null hypotheses were tested for each item by using the F statistic. Table 6, pages 50 and 51, reports the computed F values.

In testing the 60 items, null hypotheses were rejected in 67 cases and retained in 353 cases. The first hypothesis, stating that there is no significant difference among the mean scores of the five states, was rejected in 25 of the 60 items. The second hypothesis, stating that there is no significant difference between the mean scores of administrators and instructors was rejected in 19 cases. The third hypothesis, stating that there is no significant difference between multi-campus and single-campus institutions was rejected in four cases. The fourth hypothesis, stating that there are no significant

Table 6. Computed F-test Results for Each of the Factors Tested

Question Number	Factor						
	State	Personnel	Campus	SxC	SxP	CxP	SxPxC***
1	3.76*	4.39*	3.87	.36	1.74	1.78	.76
2	1.84	1.99	8.33*	.64	1.08	.01	.42
3	1.46	.35	.18	.60	1.21	.01	2.55*
4	.37	.70	.70	1.41	.15	1.66	.54
5	.21	.67	.87	1.44	1.32	.00**	1.31
6	2.84*	.97	2.91	4.36*	.32	1.52	.72
7	4.77*	11.35*	2.84	3.36*	1.67	6.06*	2.91*
8	1.09	3.73	.11	.85	.66	.75	1.01
9	4.54*	6.15*	.34	.70	1.03	3.60	.98
10	6.81*	.06	.02	.21	1.93	.26	.82
11	2.51*	5.61*	.77	1.64	1.08	3.34	.82
12	2.59*	.15	.02	3.73*	.43	1.40	2.00
13	1.25	.14	1.22	.35	.22	.65	1.71
14	18.89*	.32	.05	.96	.56	.32	.30
15	8.63*	4.86*	4.86*	.24	1.40	.60	.76
16	.67	1.07	.84	.35	.59	.03	.18
17	6.70*	.39	.01	3.23*	.24	1.08	.19
18	1.99	.40	.10	1.99	1.10	.62	.83
19	.36	8.13*	.18	.70	.62	.08	4.69*
20	2.01	.37	2.02	.21	.60	.04	.52
21	4.10*	5.03*	.31	2.14	1.60	.05	.58
22	2.04	.16	2.09	.69	.40	.03	.70
23	6.12*	12.78*	.44	2.50*	2.57*	3.09	.61
24	3.20*	.14	.98	1.36	2.57*	.02	.51
25	1.39	3.63	.77	.38	.34	.19	.90
26	4.77*	.07	.40	.53	.48	1.00	.26
27	.25	3.24	.25	2.35	1.22	.00**	.91
28	2.28	.12	0	.20	.38	.46	1.12
29	7.77*	.06	.57	4.47*	.19	.57	.69
30	1.34	8.87*	.00**	.50	.51	1.16	.53

\*Indicates significant difference at the .05 level.

\*\*A non-zero value rounded to the nearest hundredth.

\*\*\*Indicates interaction of factors: SxC represents interaction of state and campus factors; SxP represents interaction of state and personnel factors; CxP represents interaction of campus and personnel factors; and SxPxC represents interaction of state, personnel, and campus factors.

Table 6. Continued

Question Number	State	Personnel	Campus	Factor			
				SxC	SxP	CxP	SxPxC
31	.83	7.38*	.23	.42	1.00	2.56	.96
32	2.59*	4.15*	1.28	.80	.05	.46	2.03
33	1.50	8.30*	.02	.64	.75	.52	.05
34	2.72*	.45	.45	.72	.81	.07	.28
35	1.11	3.15	.04	1.10	1.38	.52	.78
36	.74	.11	1.69	2.20	.98	4.46*	.45
37	3.22*	1.98	.10	.66	1.11	2.95	2.05
38	1.71	1.18	.10	.73	.52	.10	.97
39	.41	5.69*	.02	.88	.57	.14	.77
40	4.82*	1.37	2.32	.88	.34	.12	.47
41	3.20*	1.39	1.79	1.12	1.59	.50	3.64*
42	1.07	1.22	.31	2.22	.22	.02	.26
43	1.57	3.82	2.40	2.18	.20	2.00	2.72*
44	1.03	8.12*	2.03	1.15	1.55	1.20	2.29
45	1.57	8.98*	.38	2.56*	.32	2.77	.72
46	2.80*	.07	.60	1.39	.21	1.42	.91
47	11.02*	.21	5.33*	1.47	1.30	1.64	.41
48	4.04*	.28	2.48	5.98*	.25	1.26	1.12
49	3.17*	7.67*	.53	.94	.43	.08	1.96
50	1.16	.04	.40	.51	.29	.04	1.05
51	2.53*	3.18	.40	.88	.23	2.74	.28
52	.84	.07	.67	.71	1.39	1.85	.60
53	1.04	.95	0	.82	.75	3.81	.54
54	1.90	.38	.14	.57	3.06*	.24	.22
55	1.22	1.64	.03	.88	1.63	2.72	.14
56	.87	4.31*	.17	2.41	1.47	.08	.30
57	2.30	2.47	.56	.28	2.91*	.04	.14
58	.42	5.47*	.39	.16	1.11	.39	.59
59	.29	14.84*	.14	.24	.76	.05	1.92
60	.77	1.43	5.71*	2.85*	1.42	.06	.85

interaction differences for the mean scores of the variables, was rejected in 19 cases. The fact that null hypotheses were rejected in only 67 cases indicates that the groups were alike in many of their responses but that they were significantly different for approximately 15 percent of the tests.

Table 7, pages 53, 54, and 55, reports the means for the states, campus, and personnel factors. When the means of all 60 items were ranked, the upper quartile included means larger than 3.21 and the lower quartile included means smaller than 2.03.

The Least Significant Difference (L. S. D.) Test was used to determine where specific differences existed between mean scores of the five states in the 25 cases where the F statistic rejected the null hypothesis. The L. S. D. Test is used to test the hypothesis that two population means are equal. Test procedures are shown in Appendix J.

Table 8, page 56, shows results of the L. S. D. tests. On six items, differences between the ranked mean scores exceeded the computed L. S. D. value and interpretations were made. On each of the remaining 19 items differences between the ranked mean scores did not exceed the computed L. S. D. value even though the computed F value exceeded the tabular value and the item was rejected. (See Table 6, page 50.) This difference may have occurred because of the wide range of scores possible when utilizing five means. For

Table 7. Means for the States, Personnel, and Campus Factors

Item No.	Overall Mean	FL	IA	State Means			Personnel		Campus Type	
				NY	TX	WA	Admin.	Instr.	Multi	Single
1	2.62	2.41	2.28	3.00	2.25	3.16	2.82	2.41	2.42	2.81
2	3.06	3.16	3.28	2.56	3.16	3.12	3.19	2.92	2.79	3.32
3	1.42	1.38	1.75	1.25	1.44	1.28	1.38	1.46	1.45	1.39
4	4.21	4.22	4.34	4.22	4.22	4.03	4.28	4.14	4.28	4.14
5	3.01	2.88	3.12	3.00	3.12	2.91	2.91	3.10	2.90	3.11
6	1.24	1.09	1.53	1.06	1.19	1.31	1.19	1.29	1.16	1.31
7	2.40	2.47	2.78	1.59	2.47	2.69	2.72	2.08	2.24	2.56
8	3.48	3.78	3.25	3.56	3.53	3.28	3.66	3.30	3.51	3.45
9	3.90	3.88	4.41	3.31	4.12	3.78	4.11	3.69	3.85	3.95
10	2.84	2.72	2.25	3.78	2.88	2.56	2.81	2.86	2.85	2.82
11	3.27	3.38	3.62	2.81	3.47	3.06	3.49	3.05	3.19	3.35
12	2.09	2.09	1.75	2.66	2.00	1.94	2.05	2.12	2.08	2.10
13	1.94	1.72	2.19	1.75	2.12	1.94	1.98	1.91	1.85	2.04
14	3.18	3.47	1.53	4.19	2.75	3.94	3.24	3.11	3.15	3.20
15	2.61	3.69	1.91	2.84	1.91	2.72	2.36	2.86	2.86	2.36
16	2.39	2.50	2.62	2.12	2.38	2.34	2.29	2.50	2.49	2.30
17	3.06	3.60	3.03	2.09	3.84	2.75	3.14	2.99	3.05	3.08
18	1.59	1.97	1.56	1.62	1.50	1.28	1.64	1.54	1.56	1.61
19	1.89	2.00	1.81	1.78	2.03	1.81	2.14	1.64	1.92	1.85
20	1.64	1.25	2.06	1.56	1.62	1.72	1.59	1.70	1.51	1.78

Table 7. Continued

Item No.	Overall Mean	FL	IA	State Means			Personnel		Campus Type	
				NY	TX	WA	Admin.	Instr.	Multi	Single
21	2.61	2.25	2.50	3.50	2.44	2.38	2.36	2.86	2.68	2.55
22	2.21	1.69	2.44	2.09	2.06	2.56	2.21	2.12	2.01	2.32
23	2.74	2.69	3.44	2.00	3.16	2.44	3.11	2.38	2.68	2.81
24	2.75	2.59	3.16	2.12	2.97	2.91	2.79	2.71	2.85	2.65
25	3.85	4.06	3.56	3.97	4.00	3.66	4.01	3.69	3.92	3.78
26	1.47	1.12	1.44	2.00	1.53	1.25	1.45	1.49	1.42	1.51
27	2.52	2.56	2.56	2.47	2.62	2.38	2.36	2.68	2.48	2.56
28	1.60	1.44	1.41	2.00	1.69	1.47	1.62	1.58	1.60	1.60
29	2.88	3.16	2.31	3.50	2.84	2.59	2.86	2.90	2.82	2.94
30	3.22	3.09	3.34	3.50	3.31	2.84	3.51	2.92	3.22	3.21
31	4.12	4.41	4.06	4.03	4.03	4.09	4.34	3.91	4.16	4.09
32	3.34	3.56	3.12	3.72	3.00	3.31	3.51	3.18	3.44	3.25
33	3.70	4.09	3.62	3.59	3.72	3.47	3.95	3.45	3.71	3.69
34	2.20	2.75	2.03	1.88	2.31	2.03	2.14	2.26	2.26	2.14
35	2.99	3.12	2.75	3.19	3.16	2.75	3.16	2.82	3.01	2.98
36	3.76	3.97	3.81	3.72	3.75	3.56	3.79	3.74	3.86	3.66
37	3.62	4.03	3.75	3.28	3.75	3.31	3.74	3.51	3.65	3.60
38	1.74	1.78	1.84	1.41	2.03	1.62	1.82	1.65	1.76	1.71
39	2.01	2.00	2.00	1.94	2.25	1.88	2.25	1.78	2.02	2.00
40	2.78	3.16	2.06	3.41	2.59	2.66	2.90	2.65	2.94	2.61

Table 7. Continued

Item No.	Overall Mean	FL	IA	State Means			Personnel		Campus Type	
				NY	TX	WA	Admin.	Instr.	Multi	Single
41	1.51	1.66	1.78	1.06	1.28	1.75	1.60	1.41	1.40	1.61
42	1.64	1.62	1.62	1.88	1.75	1.31	1.74	1.54	1.69	1.59
43	2.02	1.97	2.38	1.69	2.16	1.91	2.20	1.84	1.88	2.16
44	3.24	3.56	2.84	3.22	3.31	3.25	3.56	2.91	3.08	3.40
45	3.37	3.97	3.06	3.44	3.09	3.28	3.76	2.98	3.29	3.45
46	2.14	1.75	1.62	2.72	2.25	2.38	2.11	2.18	2.05	2.24
47	2.88	2.38	1.78	3.75	2.62	3.88	2.94	2.82	2.60	3.16
48	1.84	1.78	1.56	1.50	2.44	1.94	1.89	1.80	1.71	1.98
49	2.18	2.50	1.78	1.97	2.56	2.06	2.41	1.94	2.24	2.11
50	3.05	2.88	2.66	3.38	3.12	3.22	3.02	3.08	2.98	3.12
51	2.48	2.28	2.28	2.75	2.12	2.94	2.65	2.30	2.54	2.41
52	1.78	1.72	1.72	1.88	2.03	1.53	1.80	1.75	1.85	1.70
53	2.40	2.41	2.16	2.75	2.47	2.22	2.50	2.30	2.40	2.40
54	2.85	3.03	2.47	3.25	2.88	2.62	2.91	2.79	2.89	2.81
55	3.07	3.06	3.50	2.97	2.97	2.84	3.20	2.94	3.05	3.09
56	3.18	3.06	3.38	2.97	3.38	3.09	3.36	2.99	3.14	3.21
57	3.08	3.06	3.50	2.62	3.09	3.12	3.22	2.94	3.01	3.15
58	3.66	3.84	3.56	3.56	3.66	3.69	3.85	3.48	3.61	3.71
59	3.48	3.47	3.34	3.56	3.59	3.44	3.80	3.16	3.45	3.51
60	2.92	3.00	2.88	3.03	3.12	2.59	3.05	2.80	3.18	2.68



Table 8. Results of Least Significant Difference Tests\*

Item No.	<u>Differences Between Means</u>									
	FL-IA	FL-NY	FL-TX	FL-WA	IA-NY	IA-TX	IA-WA	NY-TX	NY-WA	TX-WA
1	no	no				no			no	
6		no	no				no			no
7			no	no			no	yes		
9			no	no		no			no	
10			no	no			no	yes		
11			no	no		no			no	
12		no	no				no			no
14			yes	no		yes			no	
15		yes				no	yes		no	
17	no		no				no		no	
21				no	yes	no				no
23		no	no			no			no	
24		no		no		no				no
26				no		no	no	yes		
29		no	no				no			no
32		no				no	no		no	
34			no			no	no		no	
37	no					no			no	no
40		no		no		no				no
41			no	no			no	no		
46	no		no						no	no
47	no		no					no	no	
48	no			no	no					no
49			no	no	no				no	
51	no	no				no			no	

\*Blank spaces have not been compared directly. "Yes" indicates a significant difference at the .05 level, and "no" indicates no significant difference at the .05 level.

the five means, the differences between the top ranked mean and the fifth ranked mean were great; however, differences between ranked individual means were, in many instances, not large enough to reach rejection levels for the Least Significant Difference tests. For the present study the L. S. D. test was restricted in its multiple comparisons capability because of the limitation which utilized only adjacent ranked mean comparisons. For the present research, no 'a priori' hypotheses were advanced with regard to the individual mean comparisons. Therefore, many of the F tests showed a rejection of the null hypothesis, but the L. S. D. test showed no such rejection.

### Summary of Means and Hypothesis Tests

#### Means

Who does appraisals? Findings indicated that the instructor's immediate supervisor provided the most important input [4.21].<sup>1</sup> Of less importance were inputs by students [3.06], and instructor self-evaluations [2.40]. Indications were that inputs by the personnel department [1.42] and by consultants and others from outside the college [1.24] were of little importance.

---

<sup>1</sup> Overall mean value computed from responses to a Likert-type five-point scale.

Purposes for appraisal. The most important purpose for conducting appraisals was reported to be developmental reasons (improvement of instruction) [3.90]. Other purposes included: encourage the retention of effective personnel [3.48], to articulate goals and objectives which affect supervisor and instructor [3.27], and for tenure decisions [3.18]. Purposes indicated to be of lesser importance included: guide to job changes (for decisions about promotions, transfers, or discharges) [2.84], to meet legal requirements established by the state [2.61], an administrator's system for generating data to demonstrate effective selection and assignment of personnel resources [2.39], wage and salary (size and frequency of pay raises or denial of raises) [2.09], and to validate personnel programs (check the success of inservice) [1.94].

Methods used. Rating scales were reported to be the most extensively used method for evaluating instructor performance [3.06]. Other methods were management by objectives (appraisal by results) [2.74], free-form essay [2.61], and group appraisal [2.21]. According to the data, less extensively used methods included: critical incidents [1.89], field review [1.64], and instructor comparisons by ranking or forced-distribution methods [1.59].

What was evaluated. The most emphasized criteria for measuring instructor performance was reported to be classroom interaction with students [4.12]. Other items which responses indicated

instructors must be skilled in to receive superior or complimentary appraisal reports included: attitude, judgement, initiative, leadership [3.85]; classroom management [3.76]; instructional planning or preparation [3.70]; commitment to institutional goals [3.62]; interaction with faculty [3.34]; and improvement since last appraisal [3.22]. Criteria of somewhat lesser importance included: contribution to the teaching profession [2.99]; participation on college committees [2.88]; objectives identified in a management by objectives plan [2.75]; participation in community functions, public relations, citizenship, involvement in service organizations [2.52]; and personal conduct in private life, moral standards [2.20]. Lower ranked items were: performance of students on standardized tests [1.74], research [1.60], and publications [1.47].

The next section of the questionnaire pertained to procedural aspects of the appraisal process and included items which were intended to elicit information to supplement the findings of the previous sections. Responses to this section indicated that: instructors generally receive a written copy of their final evaluation report [3.37] and that performance appraisal procedures usually are explained to new instructors [3.24]. It was also indicated that the following statements describe appraisal procedures seldom used in the colleges sampled: students and other non-professionals receive instruction to prepare them for their role in the appraisal process [1.84], salary

increments are withheld from instructors whose performance appraisal indicates unsatisfactory performance [1.64], and an instructor has the option to specify who will perform the appraisal of his performance [1.51].

Responses to items in the remainder of the questionnaire indicated that appraisals have been reasonably just and fair [3.66] and that colleague involvement in the appraisal process was not seen as a threat to the dean's position as instructional leader of his staff.

In brief, instructor performance is appraised by the instructor's immediate supervisor in most cases. Lesser inputs are provided by other administrators, students, and others. Improvement of instruction was indicated as an important purpose for performing appraisals. Encouraging retention of effective personnel, articulation of goals and objectives which affect the supervisor and instructor, and collecting information needed for tenure decisions were also indicated as purposes for conducting appraisals. The data indicated that rating scales were the most prevalent method used for collecting appraisal information. Other aspects of instructor performance appraisal, in descending order of importance included: classroom interaction with students; attitude, judgement, initiative, leadership; instructional planning; and improvement since last appraisal. In every case these generalizations represent an averaging of responses that involved less than total agreement.

### Hypothesis Testing

States. The null hypothesis, stating that there is no significant difference among the mean scores of the five states included in the study, was rejected in 25 instances. Three-way Analysis of Variance tests ( $\alpha = .05$ ) indicated that significant differences did occur in the mean scores of the five states for those 25 items. A Least Significant Difference Test on each of these items indicated that in six instances significant differences occurred between adjacent ranked means. In the remaining items the differences were evenly distributed, and significant differences among the adjacent ranked means did not occur.

Items for which the null hypothesis was rejected included:

1. Importance of formal and informal inputs made by: (1)<sup>2</sup> colleagues, (6) consultants and others from outside the college, and (7) the instructor (self-evaluation) [New York, smallest].<sup>3</sup>
2. Importance attached to conducting appraisals for: (9) developmental reasons, (10) a guide to job changes

---

<sup>2</sup>Number of the item on the questionnaire: see Appendix A.

<sup>3</sup>Bracketed statements indicate states in which the mean of the responses was significantly different (larger or smaller) than adjacent ranked means of the other states, as determined by the L. S. D. test.

[New York, largest], (11) understanding, (12) wage and salary, (14) tenure decisions [Iowa, smallest; Texas smaller than the other three], and (15) meeting state legal requirements [Iowa and Texas were identical and smaller than the other three; Florida, largest].

3. Extent to which instructor performance is appraised by: (17) rating scales, (21) free-form essay [New York, largest], and (23) appraisal by results, management by objectives.

4. Amount of emphasis placed on: (24) objectives identified in a management by objectives plan, (26) publications [New York, largest], (29) participation on college committees, (32) interaction with faculty, (34) personal conduct in private life, and (37) commitment to institutional goals.

5. Extent to which (40) teaching performance is appraised by the same evaluation(s) which provides information for decisions concerning increments, assignments, and/or tenure, (41) an instructor has the option to specify who will perform the appraisal of his performance, (46) the performance of instructors with tenure or with three or more years of experience at your college is evaluated on alternate years or even less frequently, (47) instructors who have been with the college three years or less are appraised more frequently than instructors with tenure or several years of acceptable service,

(48) students and other non-professionals receive instruction to prepare them for their role in the appraisal process, (49) evaluators are adequately trained in such data gathering techniques as classroom interaction analysis and task surveys, and (51) colleague evaluations are supported by instructors.

New York responses showed the greatest number of variations. The data indicated that in New York: self-evaluations by the instructor are less important, appraisals as a guide to job changes are more important, free-form essay evaluations are used more extensively, and more emphasis is placed on publication than in the other four states. Iowa and Texas responses indicated that tenure decisions and meeting state requirements tend to be less important reasons for conducting appraisals. Florida responses indicated the most importance attached to appraisal for meeting state legal requirements. In general, the states tended to be more alike than different, but responses indicated significant differences in 25 of the 60 items on the questionnaire.

Personnel. The null hypothesis, stating that there is no significant difference between the mean scores of administrators and instructors, was rejected in 19 instances. In 17 of those instances the mean of responses by administrators was larger than the mean of responses by instructors.



Items for which mean responses by administrators was larger included:

1. Importance of formal inputs made by (1) colleagues and (2) the instructor (self-evaluation).

2. Importance attached to (9) developmental reasons and (11) understanding.

3. Amount of emphasis placed on (19) critical incidents and (23) appraisal by results, management by objectives.

4. Amount of emphasis placed on: (30) improvement since last appraisal, (31) classroom interaction with students, (32) interaction with faculty, and (33) instructional planning or preparation.

5. Extent to which: (39) evaluative criteria (standards) are developed for each teaching position, (44) performance appraisal procedures are explained to all new instructors, (45) instructors receive a written copy of their final evaluation report, (49) evaluators are adequately trained in such data gathering techniques as classroom interaction and task surveys, (56) instructor performance appraisal procedures lead to improved classroom instruction, and (59) most evaluators have an adequate understanding of teaching methods.

Items for which the mean of responses by instructors was larger include:

1. Importance attached to conducting appraisals for (15) meeting legal requirements established by the state.

2. Extent to which instructor performance is appraised by (21) free-form essay.

In 17 of the 19 items where significant differences occurred, the mean for administrator's responses were larger than the mean for instructor's responses. Significant differences occurred in approximately one-third of the items on the questionnaire.

Campus. The null hypothesis, stating that there is no significant difference between the mean scores of multi-campus and single-campus institutions, was rejected in four instances. Responses from multi-campus colleges indicated (15) greater importance attached to conducting appraisals to meet legal requirements established by the state, (47) greater agreement with the statement that instructors who have been with the college three years or less are appraised more frequently than instructors with tenure or several years of acceptable service, and (60) greater agreement with the statement that given a choice between a colleague or an administrative evaluation, most instructors would select an administrative evaluation. Responses from single-campus colleges indicated that (2) more importance is attached to inputs made by students.

Interaction. The null hypothesis, stating that there are no significant interaction differences for the mean scores of the variables

included in the study, was rejected in 23 instances. Nine of those instances involved interaction of the state and campus factors, four involved interaction of the state and personnel factors, two involved interaction of the campus and personnel factors, and eight involved the combined interaction of state, personnel, and campus factors.

#### Data Describing Classroom Observations

The data in part H of the questionnaire were provided only by personnel in colleges where classroom visitations were included in the appraisal process. Responses to these items indicated that approximately 58 percent of the colleges included classroom observations in their appraisal procedures. Twenty-two percent of the 160 responses indicated one observation per appraisal, 18 percent indicated two or three observations, five percent indicated four or more observations, and 13 percent indicated that the number of observations per appraisal varies with instructors.

Those reporting classroom visits indicated that the duration of visits in over 56 percent of the colleges was over 30 minutes. Seven percent of the responses indicated that classroom observations were 10 minutes or less in length, 20 percent indicated that observations were 11 to 20 minutes in length, and 17 percent indicated that observations were 21 to 30 minutes in length.

Of the responses indicating that classroom visitations were made, 15 percent indicated that no post-observation conferences were held. Twelve percent indicated that post-observation conferences were held the same week, and 24 percent indicated that conferences take place more than a week after the observation.

The next item on the questionnaire inquired about the helpfulness of post-observation conferences. Of those responding to the question (a total of 85) forty-seven percent indicated that post-observation conferences were very helpful, 49 percent indicated that these conferences were sometimes helpful, and four percent indicated that they were seldom helpful.

The final item in this section provided an opportunity for responses about a variety of procedures related to classroom observations. The percentages cited refer only to persons who indicated that classroom observations are made at their colleges. Eight percent of the responses indicated that audio tapes were made of the observations, and 11 percent indicated that video tapes were made. Seven percent indicated that audio-video tapes were reviewed during the post-observation conferences. Fifty-eight percent indicated that written summaries of the observations were presented to the instructor after each observation; 29 percent reported the use of pre-observation conferences; and 24 percent indicated that evaluators make observations only when invited by the instructor.

In summary, approximately 58 per cent of the responses indicated that classroom observations were a part of the instructor appraisal process. The majority reported that observations exceed 30 minutes in length. Nearly half of the responses indicated that post-observation conferences are held within a week of the observation, and the responses indicated that the conferences are helpful. More than half of the responses indicated that written summaries of classroom observations were presented to the instructor.

## CHAPTER V

CONCLUSIONS, IMPLICATIONS, AND SUGGESTIONS  
FOR FURTHER STUDY

This study was designed to determine performance appraisal methods being used in community colleges (in Florida, Iowa, New York, Texas, and Washington) and to identify factors that accounted for variations in performance appraisal procedures. Questions about instructor performance appraisal which were examined included: who does it, for what purposes, what methods are used, and what is evaluated. Mailed questionnaires elicited responses from community college administrators and instructors in multi-campus and single-campus colleges. Data obtained from the questionnaires were used to test null hypotheses which stated that significant differences do not occur (1) among the states, (2) between administrator's responses and instructor's responses, (3) between the responses from multi-campus colleges and single-campus colleges, and (4) among the various combinations of state, personnel, and campus factors.

ConclusionsWho Does Appraisals and Why

The data indicated that the instructor's immediate supervisor provides the most important input in the appraisal

process.<sup>4</sup> A review of the purposes for conducting appraisals may reveal certain dysfunctional aspects of heavy reliance on evaluations by the immediate supervisor.

The three most important reasons for conducting appraisals were reported to be: developmental reasons (improvement of instruction), encouraging retention of effective personnel, and understanding (to articulate goals and objectives which affect the supervisor and instructor). These create conflict for the supervisor.

A review of the data indicates that:

1. Only 58 percent of the responses indicated classroom observations are made. Yet, instruction is to be improved, "understanding" increased, and the evaluator presumably must monitor the improvement.

2. Appraisal is to encourage the retention of effective personnel and to lead to greater understanding, but it seems unlikely that instructors will reveal inadequacies which could later be used for management decisions such as promotion or tenure.

---

<sup>4</sup>When the means of all 60 items are ranked, 25 percent of the means are larger than 3.21 and 25 percent are smaller than 2.03. Much of the discussion in this chapter will center around items in the upper quartile range.

Accomplishment of these purposes for appraisals would seem more likely to occur if the supervisor were responsible primarily for supervision of instruction. Much of the information needed for management decisions (promotion, tenure, and other rewards) could be obtained from evaluations by students, colleagues, and other administrators.

#### Methods of Conducting Appraisals

The means for items in this section of the questionnaire were all in the lower three quartiles, indicating that none of the methods listed were extensively used throughout the population sampled. The fact that the questionnaire was completed provides an indication that appraisals were made. Reasons for lack of large means in this section may include:

1. Inconsistencies within the colleges: more than one method or a combination of methods may be used.
2. Respondents may have not recognized terms and descriptions which were applicable to the appraisal process in their colleges.
3. Methods not listed and distinctly different from those on the questionnaire may be used.

Administrators indicated more extensive utilization of: (1) critical incidents and (2) appraisal by results (management by objectives).



Instructors indicated greater use of free-form essay. These conflicts suggest that:

1. Appraisal procedures are not as thoroughly explained as indicated in item 44 (performance appraisal procedures are explained to all new instructors),
2. Changes in appraisal procedures have not been explained as the process has developed, or
3. Administrators, because of their proximity with the appraisal process have a perspective which differs from the perspective of instructors.

Even though one method was not indicated to be prevalent, the difference between administrator's responses and instructor's responses indicates a lack of effective communication about appraisal methods.

#### Importance of Items Which Were Appraised

The responses indicated a greater emphasis on the type of behavior appraised than on the method used. A total of seven items occurred in the upper quartile when the means for the items on the questionnaire were ranked.

Classroom interaction with students was reported to receive the greatest emphasis. Classroom management was also reported to be emphasized. In view of the fact that only 58 percent of the

respondents indicated use of classroom observations, it would appear either that (1) information about classroom performance was obtained from students or that (2) some evaluations involved judgements not based on direct observations.

### State Differences

Responses from the state of New York were more frequently significantly different than the responses from any other state. Two characteristics of New York community colleges which may contribute to some of these differences are: (1) greater involvement of unions in the evaluation process and (2) organizational differences which result from affiliation with the university systems in the state. All New York community colleges are attached administratively to either the State University of New York or the City University of New York, and that organizational influence may partially account for the emphasis in New York on publications by the instructor. Responses to item 10, combined with responses in the comments section, suggest that union influences may have resulted in the necessity to establish a defensible system for making decisions about promotions, transfers, or discharges.

Tenure decisions were reported to be less important in Iowa than in the other states. That finding was not surprising since many

of Iowa's newly-formed community colleges have not been in operation long enough for instructors to have earned tenure.

State laws apparently provide more impetus for appraisal of instructor performance in Florida than in the other states.

Although significant differences between responses from the five states were revealed by the F statistic, the Least Significant Difference Test was only partially successful in identifying the major source of variation.

#### Personnel Differences

Significant differences between administrator's responses and instructor's responses occurred in nearly one third of the items and the mean for administrator's responses was larger in all but two cases. Administrators tended to report: greater involvement of faculty in the appraisal process, more emphasis on improvement of instruction, and more emphasis on articulation of goals and objectives. Administrators also indicated greater amounts of communication about the appraisal process, and greater confidence in the success of the appraisal system. In general, the evaluators perceive the process of instructor appraisal more positively than do those being appraised.

### Campus Differences

Only four significant differences occurred between responses from multi-campus colleges and single-campus colleges. Responses from multi-campus colleges indicated (1) more frequent appraisal of instructors with three years or less experience, (2) appraisals are made to meet state requirements, and (3) more reliance is placed on administrative evaluations. Single-campus responses indicated greater involvement of students in the appraisal process. Multi-campus responses may indicate a somewhat greater tendency to streamline administration. Single-campus responses suggest the possibility that students are more actively involved in appraisal when the college administration is located nearby (on the same campus). Generally, multi-campus and single-campus colleges tend to be similar.

### Implications

The findings of this study and the related literature which was reviewed indicate that instructor appraisal should include:

1. A definition of organizational objectives (jointly agreed upon by students, the board of trustees, faculty, and possibly others such as advisory groups from the community).
2. Coordination between instructor and supervisor in planning a program of work which describes how institutional

objectives can be accomplished in a manner which is compatible with the personal goals of the instructor.

3. Standards of performance and a system of evaluation (planned before the program of work is implemented).

4. Collection of appraisal data as preplanned.

5. A conference with the instructor to review findings of the appraisal, followed by a written summary of the findings for the instructor's file.

6. Review and revision of the appraisal process.

The immediate supervisor's role in supervision should continue to be important, but greater involvement of other personnel in appraisal of instructor performance seems desirable. Evaluations for management purposes such as promotion or tenure decisions should involve inputs from more than one source. Administrators (including the supervisor) and colleagues should be involved in a group appraisal process, and input from student evaluations should comprise a significant portion of the data reviewed. These inputs are desirable because:

1. The additional perspective gained by including other administrators is protection for (1) the instructor against biases which the supervisor may have and (2) the supervisor against concerns of the instructor about possible biases or inadequacies.

2. Students can provide data which, interpreted by a qualified person, may provide (1) the only assessment of performance in the classroom (as indicated by 42 percent of the responses) or (2) information which can be compared with that collected during classroom observations.

The composition of the appraisal group and the precise role of each member seems to be an issue which can best be resolved as each college develops its appraisal system. Classroom visitations made for evaluations to be used for management purposes probably should be performed by somebody besides the instructor's immediate supervisor.

The immediate supervisor should make classroom observations for the purpose of assisting in the improvement of instruction. A comprehensive system of clinical supervision should be implemented, and all first-year instructors should be evaluated. Additional evaluations should be (1) performed when requested by an instructor, (2) repeated according to a schedule established as the college formulates its appraisal procedures, and (3) repeated when recommended as the result of a recommendation of an appraisal made for management purposes. The system should include:

1. A preobservation conference between the instructor and supervisor to discuss lesson objectives, teaching strategies, and the conduct of the appraisal.

2. A period of observation in which the observer(s) record data describing the instructor's teaching patterns: a system of observation such as that proposed by Goldhammer (1969).

3. Analysis of the session and planning of a strategy for presenting the observations to the instructor.

4. Supervision conference with the instructor to review observations and discuss possible modifications for improvement of teaching.

5. Post-conference analysis to critique the effectiveness of the preceding sequence.

Clinical supervision should be aimed only at improvement of instruction, and observations during these sessions should not become a part of the instructor's appraisal record. Audio/video recordings are usually helpful in reviewing the classroom observations, but use of recordings should be avoided if the instructor is apprehensive about having his performance permanently recorded. Taped sessions should be erasable at the instructor's request, and findings of clinical supervisory sessions should be confidential information. The payoff of this supervision would be improved instruction which should be reflected in the group appraisal used for management purposes.

Preservice and inservice instruction is recommended for all persons involved in the appraisal process. Specific recommendations include:

1. Each year all instructors should be briefed on the appraisal process (in group meetings) and the use of clinical supervision solely for improvement of instruction should be clarified.

2. Administrators should receive instruction to prepare them with (1) personnel management skills, (2) supervisory skills, and (3) a competency for conducting clinical supervision; inservice should be planned to keep administrators up-to-date and to prepare new administrators with skills not gained during preservice instruction.

3. Students should be informed of the purpose and importance of data gathering systems used to indicate impressions of instructor performance.

A group comprised of representatives from the faculty (administrators and instructors), students, and the board of trustees should select the method of collecting appraisal data. A faculty-designed checklist with provisions for free-form essay comments is recommended for collecting data from students.

Areas of instructor competence to be evaluated should also be selected by each college faculty. If a management by objectives



approach is used, the final decision on what is evaluated will rest with the immediate supervisor and the instructor. Otherwise, a group process designed to gain faculty consensus should be employed. The findings of this study indicate that items which many colleges consider important include: classroom interaction with students, attitude, judgement, initiative, leadership, classroom management, instructional planning or preparation, commitment to institutional goals, interaction with faculty and improvement since last appraisal.

In conclusion, administrators and instructors have a joint responsibility for planning and implementing instructor performance appraisals. The success of appraisals is dependent on clearly defined roles and thoroughly informed participants.

#### Suggestions for Further Study

The present research and the accompanying review of related literature identified questions which could provide the basis for further research. Suggestions for further study include:

1. Identify skills or competencies needed by community college administrators for performing their various roles in instructor performance appraisal.
2. Compare and evaluate effectiveness of various appraisal systems.

3. Analyze the ability of appraisal systems to provide a legal basis for personnel decisions in harmony with federal laws.

## BIBLIOGRAPHY

- Barr, A. S. 1961. Wisconsin studies of the measurement and prediction of teacher effectiveness. Madison, Dembar Publications. 156 p.
- Beach, Dale S. 1970. Personnel, the management of people at work. London, Macmillan, 844 p.
- Blocker, Clyde E., Robert H. Plummer, and Richard C. Richardson. 1965. The two-year college: a social synthesis. Englewood Cliffs, Prentice-Hall, 228 p.
- Bolin, John G. and John W. Muir. 1966. Merit rating for salary increases and promotions. Athens, Georgia University. 42 p.
- Bolton, Dale L. 1972. Teacher evaluation. In: Research Reports on Teacher Evaluation, prepared by Dale L. Bolton. Salem, Oregon Board of Education. Various paging. (Putting Research into Educational Practice. PREP 21)
- Boyer, Marcia. 1970. Teacher evaluation: toward improving instruction. Junior College Research Review, Vol. 4, No. 5, n. p.
- Burge, Ronald Dwayne. 1972. Development of a system of educator evaluation in a middle sized city school system. Ed. D. Dissertation. Eugene, University of Oregon, 215 numb. leaves.
- Caplow, Theodore and Reece McGee. 1958. The academic marketplace. New York, Basic Books, 262 p.
- Centra, John A. 1971. Problems of evaluating teaching. In: Faculty evaluation working papers: 1970-71 revision. Florida Community Junior College Research Council, p. 13-16. (Educational Resources Information Center. ED 042 934) (Microfiche)
- Cohen, Arthur M. and Florence B. Brawer. 1969. Measuring faculty performance. Monograph No. 4. American Association of Junior Colleges, Washington, D. C., 90 p.
- Court Decides, A. 1973. The Education Digest 38:68. Mar.

- Fattu, N. A. 1963. Research on teacher evaluation. National Elementary Principal 43:19-27. Nov.
- Flanders, Ned A. 1964. Interaction analysis. Ann Arbor, University of Michigan, School of Education. 45 p.
- Freund, John E. 1970. Statistics, a first course. Englewood Cliffs, Prentice-Hall. 340 p.
- Gleazer, Edmund J. Jr. 1971. American junior colleges. Meansha, Banta, 850 p.
- Goldhammer, Robert. 1969. Clinical supervision. New York, Holt, Rinehart, and Winston. 370 p.
- Hain, John H. and George J. Smith. 1968. How principals rate teachers. American School Board Journal 47:17-19. Feb.
- Harsh, Richard J. 1970. A look at teacher evaluation. Las Vegas, A paper presented at the annual conference of School Personnel Administrators, 20 p. (Educational Resources Information Center, ED 046 993) (Microfiche)
- Hayes, Robert B. 1968. A way to evaluate and to improve classroom teaching effectiveness. Harrisburg, Pennsylvania State Department of Education, 10 p.
- Howsam, Robert B. 1963. Facts and folklore. National Elementary Principal 43:7-8. Nov.
- Iowa, State Department of Public Instruction. 1972. Opportunities in Iowa's area schools, 1971-1972. Des Moines, 71 p.
- Iowa, State Department of Public Instruction. 1971. Selected school laws and standards. Des Moines, 23 p.
- Jenkins, Charles. 1970. Evaluation at Highland Community College. In: Selected papers from Northern Illinois University Community College Conferences, 1969-1970. De Kalb, Northern Illinois University, p. 1-2. (Educational Resources Information Center, ED 042 445) (Microfiche)
- Kast, Fremont E. and James E. Rosenzweig. 1970. Organization and management a systems approach. New York, McGraw-Hill, 654 p.

- Lahti, Robert E. 1970. Harper College organizational development and the implementation of a management by objectives system in an educational environment. Los Angeles, University of California. 26 p.
- Langley, Dorothy. 1972. Telephone communication concerning enrollment data in community college enrollment reports from 1959 to date. Salem, State Department of Education. Oct. 11.
- McCall, William A. 1952. Measurement of teacher merit. Publication No. 284. Raleigh, North Carolina: State Superintendent of Public Instruction. 42 p.
- McGregor, Douglas. 1960. The human side of enterprise. New York, McGraw Hill. 246 p.
- Mc Phail, James M. 1967. Teacher evaluation--a state-by-state analysis. The Southern Journal of Educational Research 1:333-356. Oct.
- Morsh, Joseph E., George G. Burgess, and Paul N. Smith. 1956. Student achievement as a measure of instructor effectiveness. Journal of Educational Psychology 2:79-80. Feb.
- Musella, Donald. 1970. Improving teacher evaluation. The Journal of Teacher Education 21:15. Spring.
- National Education Association, Research Division. 1964. Evaluation of classroom teachers. Washington, D. C. 125 p. (Research Report R14)
- Odiorne, George S. 1965. Management by objectives. New York, Pitman, 198 p.
- Olson, Claire T. 1972. The U. S. challenges discrimination against women. Junior College Journal 42:13-16. June/July.
- Oregon Board of Education. 1967. Introduction to linear estimation and hypothesis testing. Corvallis, Oregon State University. 280 p.
- Rausa, Frank. 1970. Impressions of evaluation of instructional staff. In: Selected papers from Northern Illinois University Community College Conferences, 1969-1970. De Kalb, Northern Illinois University, p. 7-10. (Educational Resources Information Center. ED 042 445) (Microfiche)

- Redfern, George B. 1963. How to appraise teacher performance. Columbus, School Management Institute, Inc., 100 p.
- Remmers, H. H. 1963. Rating methods in research on teaching. In: Handbook of Research on Teaching. Chicago, Rand McNally. p. 368.
- Rose, Gale W. 1963. The effects of administrative evaluation. National Elementary School Principals 43:50-53. Nov.
- Schafer, Michael I. 1971. Faculty evaluation working papers: 1970-71 revision. Florida Community Junior College Research Council, 130 p. (Educational Resources Information Center. ED 042 934) (Microfiche)
- Schenck, Peter. 1973. No specific sex discrimination found. Gazette-Times (Corvallis, Oregon) p. 1, col. 1-2-3-4-5. Feb. 9.
- Schultz, Raymond E. 1970. The role of the academic dean. In: Selected papers from Northern Illinois University Community College conferences, 1969-1970. De Kalb, Northern Illinois University, p. 16-24. (Educational Resources Information Center. ED 042 445) (Microfiche)
- Shapely, Deborah. 1972. University women's rights: whose feet are dragging? Science 175:151-154. Jan. 14.
- Wattenbarger, James L. 1971. Evaluation of faculty by students. In: Faculty evaluation working papers: 1970-71 revision. Florida Community Junior College Research Council, p. 5-12. (Educational Resources Information Center. ED 042 934) (Microfiche)
- Webster's seventh new collegiate dictionary. 1967. Springfield (Massachusetts), G. C. Merriam. 1219 p.
- Willard, Wayne E. 1970. Evaluation of instruction. In: Papers from Northern Illinois University Community College Conferences, 1969-1970. De Kalb, University of Northern Illinois. p. 11-16. (Educational Resources Information Center. ED 042 445) (Microfiche)

- Wisgoski, Alfred. 1970. Instructional staff evaluation. In: Selected papers from Northern Illinois University Community College Conferences, 1969-1970. De Kalb, Northern Illinois University. p. 3-6. (Educational Resources Information Center. ED 042 445) (Microfiche)
- Woodburne, Lloyd. 1950. Faculty personnel policies in higher education. New York, Harper, p. 151-193.

## APPENDICES



## QUESTIONNAIRE

Definitions:

Performance appraisal: Systematic evaluation of a full-time instructor's performance on the job and his potential for development. Other terms used to convey similar meaning include personnel appraisal, personnel review, progress report, service rating, performance evaluation, and fitness report.

Supervisor: The individual to whom an instructor reports. The individual who is most directly responsible for overseeing the work of the instructor.

Instructions:

Various parts of this questionnaire follow a programmed format; i.e., your response to the questions will determine what further action you should take. Follow the directions given after the response you check.

Part A

1. Place an X in one of the boxes below to designate whether you are an instructor or an administrator. If your work includes both types of responsibilities mark the one which consumes the greater portion of your time.

Administrator

Instructor

2. Place an X in the box preceding the statement below which most accurately describes the nature of the personnel appraisal process at your college.

The appraisal process includes formal procedures which are clearly defined and described in inservice meetings and/or documents which are readily available to instructors. (Go to # 3.)

As I understand the appraisal process it is informal and not clearly defined. My responses will describe the process as I see it. (Go to # 3.)

No appraisal process (formal or informal) is evident at our college; i.e., no apparent effort is made to evaluate quality of instruction, to determine who should receive raises, or to justify promotions, dismissals, or other personnel actions. (If this response is checked stop at this point and return the questionnaire.)

3. Is the performance of all full-time instructors at your college appraised by essentially the same system? Place an X in the box preceding the statement which describes who is appraised by the system you will be referring to as you complete this questionnaire.

The performance of all our instructors is appraised by essentially the same system.

Vocational-technical instructors only

Lower division (transfer) instructors only

Instructors in the \_\_\_\_\_ division only (Specify which one.)

Others (Specify who.) \_\_\_\_\_

## APPENDIX B

(Oregon State University Letterhead)

---

---

---

Dear \_\_\_\_\_:

Research is currently underway at Oregon State University to investigate faculty performance appraisal processes in the community colleges of Florida, Iowa, New York, Texas, and Washington. The study will analyze (1) differences in the way the appraisal process is viewed by instructors and administrators, (2) differences between states, and (3) differences in the way appraisal is handled between multi-campus and single-campus colleges. The intent of the study is to identify refinements which will permit more effective use of performance appraisal for the improvement of instruction.

We would appreciate your permission to obtain data for this study by mailing a questionnaire to administrators and instructors in your college. A total of 48 names will be selected at random from community colleges within your state. The exact number representing your institution will not be known until the sample is drawn. Completion of the questionnaire should require less than 30 minutes of the respondent's time. The individuals responding will not be identified in the final report. Copies of the report will be made available to the participating colleges.

Enclosed is a self-addressed postal card. Your cooperation in checking and returning the card at your earliest convenience will be very much appreciated. If you will participate in the study, please either send a faculty directory or indicate to whom we may write to obtain one.

Thank you for your consideration of this matter.

Sincerely,

David Maxwell, Career Education Intern  
Oregon Board of Education

## APPENDIX C

## Sample of Postal Cards Sent to Presidents

(Front Side)

	STAMP
Dave Maxwell Oregon Board of Education 942 Lancaster Drive N. E. Salem, Oregon 97310	

(Back Side)

_____	Yes, this college may be included in the 5-state faculty performance appraisal study. A copy of our faculty directory may be obtained from _____
_____	No, this college will not participate.
(Name of college was typed here prior to mailing to presidents.)	

APPENDIX D  
(Oregon State University Letterhead)

MEMO

To: \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

Research is currently underway at Oregon State University to investigate faculty performance appraisal processes in the community colleges of Florida, Iowa, New York, Texas, and Washington. The study will analyze (1) differences in the way the appraisal process is viewed by instructors and administrators, (2) differences among states, and (3) differences in the way appraisal is handled between multi-campus and single-campus colleges. The intent of the study is to identify refinements which will permit more effective use of performance appraisal for the improvement of instruction.

The President of your college has consented to include the college in this study and has indicated that you can provide a directory of faculty personnel. The persons listed in the directory will be included in the group from which a sample will be drawn. The directory will be used only for this study and will not be released for other purposes.

Please send a copy of the faculty directory at your earliest convenience. A self-addressed label is enclosed for your convenience in mailing the directory. If any charges for postage and handling are involved please indicate to whom reimbursement should be submitted. The cooperation of your college is cordially appreciated.

Sincerely,

David Maxwell, Career Education Intern  
Oregon Board of Education

Enclosure

(Oregon State University Letterhead)

\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

Dear \_\_\_\_\_:

Research is currently underway at Oregon State University to investigate faculty performance appraisal processes in the community colleges of Florida, Iowa, New York, Texas, and Washington. The study will analyze (1) differences in the way the appraisal process is viewed by instructors and administrators, (2) differences among states, and (3) differences in the way appraisal is handled between multi-campus and single-campus colleges. The intent of the study is to identify refinements which will permit more effective use of performance appraisal for the improvement of instruction.

A letter was mailed before Christmas asking for permission to include your college in this study. We are still anxious to have your college participate but have not received the card which was to have been returned indicating if you prefer to be included or excluded. If you have returned the card please consider this a letter of appreciation. We would appreciate your checking and returning the enclosed card, just in case the previous card was misplaced or delayed in the mail.

If we may include your college in the study please send a directory of faculty personnel so your faculty members may be included in the list from which a random sample will be drawn. The directory will be used only for this study and will not be released for other purposes. We expect that less than six persons will be selected from any one college. Individuals participating in the study will not be identified in the final report. Copies of the report will be made available to the participating colleges.

Thank you for your consideration of this matter.

Sincerely,

David Maxwell, Career Education Intern  
Oregon Board of Education

Enclosure

## APPENDIX F

(Oregon State University Letterhead)

---

---

---

Dear \_\_\_\_\_:

Research is currently underway at Oregon State University to investigate faculty performance appraisal processes in the community colleges of Florida, Iowa, New York, Texas, and Washington. The study will analyze (1) differences in the way the appraisal process is viewed by instructors and administrators, (2) differences among states, and (3) differences in the way appraisal is handled between multi-campus and single-campus colleges. The intent of the study is to identify refinements which will permit more effective use of performance appraisal for the improvement of instruction.

Your President has indicated his willingness to permit us to include personnel at your college in the study. This is not a study of individuals or of individual colleges, and respondents will not be identified in the final report. A summary of the study will be mailed to the president's office in participating colleges. Provision is made on the questionnaire for you to indicate your desire to receive a personal copy.

Please complete and return the questionnaire which accompanies this letter. (In a pilot survey the average time required to answer the questions was less than 20 minutes.) Your professional assistance will be cordially appreciated.

Sincerely,

David Maxwell, Career Education Intern  
Oregon Board of Education

Enclosure

(Oregon State University Letterhead)

\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

Dear \_\_\_\_\_:

With the approval of your president, I recently mailed a study questionnaire requesting your help in obtaining data for a study of instructor performance appraisal procedures in community colleges. If you have already completed and returned the questionnaire, please consider this note as an expression of our appreciation.

If you have not responded please complete and return the questionnaire by March 30. Another copy of the questionnaire is enclosed for your convenience.

Thank you for your cooperation.

Sincerely,

David Maxwell, Career Education Intern  
Oregon Board of Education

Enclosure

APPENDIX H  
Coding of Data Cards

Column	
1	1 for card number one (first and only card for each questionnaire).
2, 3, 4	Three digit numbers from 001 through 160 representing numbers which identify respondents.
5	State number: 1 for Florida, 2 for Iowa, 3 for New York, 4 for Texas, and 5 for Washington.
6	Campus type: 1 for multi-campus and 2 for single-campus.
7	Personnel type: 1 for administrator and 2 for instructor.
8	Part A, item 2: 1 if first box was checked and 2 if second box was checked.
9	Part A, item 3: 1 if first box was checked, 2 if second box was checked, 3 if third box was checked, 4 if fourth box was checked, and 5 if fifth box was checked.
10	Part A, item 4: 1 if "colleagues" was checked or a 0 if not checked.
11 through 17	Part A, item 4: 1 if the item was checked and a 0 if the item was not checked.
18 through 77	Parts B, C, D, E, F, and G; items 1-60: digits 1 through 5 representing the value marked by the respondent on the five-point interval scale



## Comments by Respondents

A space was provided at the end of the questionnaire for comments. The following comments were taken directly from the questionnaires.

Florida Administrators

These are my judgments only: tenured faculty vote on reappointment and tenure. I do not believe it is possible at present to measure teaching effectiveness. All methods are fraught with problems.

We are not sure that our evaluation separates the instructors by quality. The spread between the high instructor and low one is very small. Generally, we are not enthused with our method.

61-66 is worked out in an informal way with an evaluating team of 3 other instructors selected by the one being evaluated. There is no established policy on classroom visits yet it is used often.

No required classroom visits as to number and length. This does take place. Some use is being made of video tapes-- on a voluntary basis.

Florida Instructors

General instructor consensus is that performance appraisal is made with little or no classroom visitation and more on personality outside the classroom.

Not enough actual knowledge of administrative procedures! Evaluations have been administered differently each year in response to faculty dissatisfaction. Evaluation ruffles feathers but doesn't seem to effect change in personnel or behavior. No one considers himself less than outstanding!

It would seem to me that the best evaluation of teaching would involve a judicious examination of student opinions, plus a system of departmental comprehensive final exams which

would show how well one instructor's students do in relation to other students and to the departmentally agreed standard exam.

This was very difficult to fill out. There is no "system" here, to speak of, and no one is informed about it. That is, the instructors do not know what they are evaluated on, and they do not see whatever evaluation is arrived at. There is also no follow-up, no attempt to use the evaluation to improve instruction. Besides, the evaluators do not visit classes or have any concrete means by which to evaluate.

Refer to item 63 above. This will vary with observer. Sometimes post-observation conference is not held.

In my opinion this is one of the best questionnaires I've seen. Good luck on your efforts. I would be interested in the results.

#### Iowa Administrators

There is no way some of these questions can be answered correctly under this rating scale.

Our evaluation program is based on each unit of the college; e. g., depts., dept. heads, deans formulating goals for the unit and evaluating their achievement by a group.

I am very much interested in the results regarding this questionnaire and your final appraisal of the study. Thank you.

I'm interested in your recommendations for correcting this needed area of instructional improvement.

#### Iowa Instructors

Many questions were non-applicable due to no formal evaluations of instruction here. Our dept. is evaluated by DPI and state board standards. Evaluation is directed toward my supervisor.

These responses are made according to a program that is in the process of being initiated.

Most all of the above questions and answers are dependent upon the staff member involved. Our program is a modified type of MBO called Growth Assessment and Development. Each staff member selects a committee, writes his objectives, selectes his methods of assessment, meets his committee to finalize procedures, carries out his program, meets with his committee for final assessment. The committee consists of immediate supervisor, a colleague, and an administrator-- this is for a staff member who is an instructor. It is similar for administrators--i. e. the college president has an instructor, an administrator, a board member.

### New York Administrators

The comments by this group were brief and dealt with the purposes for conducting appraisals. One comment indicated that union contract provisions provide a very important reason, and another comment indicated that appraisals are made "to fairly reward good work performed."

### New York Instructors

The observation report by a member of an elected faculty comm. (Personnel and Budget Comm.) is shared by the chairman of the dept. and the faculty member being observed. There is no real face to face confrontation of observed and observer. This is governed by univ. by-laws.

The school relies on classroom evaluations for tenure and promotion, but it is difficult to say how much this counts. If an instructor is extremely poor--chances are he won't be rehired, but if he is very good there is no guarantee of tenure or promotion.

The majority of the faculty in a poll have shown support of seniority as a basis for promotion--rather than the subjective political method of attempting to rate according to "teaching effectiveness." A solid probationary period is the safeguard! After that--obvious deterioration of teaching could be the main cause of severance.

### Texas Administrators

Answers may be a bit misleading in that some of questions do not exactly fit situation at this institution.

College has area Coordinators directly over instructor and Division Director over Coordinator. Both can evaluate. Different Division Directors evaluate differently. I always do classroom observation at least once per year, but some Directors rarely do. Part of our philosophy concerns evaluation. Instructors are informed of the process in their initial interview.

We have no tenure law.

Multi-campus--as a district admin., this questionnaire is difficult to answer. Each campus has its own system of faculty evaluation.

Reacting to item 42 (Salary increments are withheld from instructors whose performance appraisal indicates unsatisfactory performance) -- "Will dismiss this type of person"

### Texas Instructors

1. Questions too involved

2. Some evaluation techniques may or may not be there. Generally faculty misinformed or ignorant.

Some of these questions are difficult to answer since we are under a new administration in its 2nd year. --also from having resigned as Division Chairman and returned to the teaching staff. The questionnaire is a good one.

#38 We have recently discussed the use of such tests.

#45 We see student evaluations but not the division director's report to the administration.

Our ultimate evaluation is: if companies return to our department and ask for more students like the ones they have

received in the past. Then we evaluate that we are doing a good job. (On the average students have six job offers when they graduate.)

We try in the technical drafting, technical math, and home building departments to create an atmosphere in which the instructor can bring problems to the department head without evaluation of the instructor's competency being questioned. I feel formal evaluations tend to force instructors to try to hide weaknesses and problems. I have instructors present when I write out evaluations and have them help me to formulate such evaluations.

### Washington Administrators

I am Dept. Chairman of the Wood Construction Dept. I do not believe this is a valid questionnaire for our vocational area. I believe you should design one that will fill our needs. Our instructors are carefully screened by their peers and related advisory committee.

Board policy mandates admin. evaluation.

### Washington Instructors

Our system is so casual and haphazard that many of my answers are of questionable validity.

Before tenure there are three classroom observations by division chairman and two faculty members, each separately. After tenure, a once-yearly observation usually takes place. This evaluation is made only by division chairman.

Most evaluation is informal except for tenure and salary increases. Informal evaluation is by Dean and Chairman and seems to center on how well the instructor maintains enrollment and holds to a not too clearly specified grading norm; i. e. not too many low grades.

Evaluation for tenure are more stringent than for other purposes. Evaluation of teaching effectiveness is aimed at providing a basis for improvement for the instructor, rather than ranking or grading.

Comments on Questionnaires Reporting "No Appraisal Process"

Program is being set up but not implemented at this date.

There is a program that has been proposed, which is to begin April 1, 1973. But as of today my answer would have to be as I have indicated.

We have discontinued our formal evaluation system pending submission of alternate plan by faculty. (This was the president's decision.)

In summary, the comments tended to indicate a lack of clearly defined appraisal procedures in some colleges, and they also indicated inapplicability of the items on the questionnaire to the appraisal procedures in some colleges. It was also indicated that the basis for tenure decisions varies from college-to-college.

## Sample Computation of L. S. D. Test

$$\begin{aligned}
 \text{L. S. D.} &= t_{\alpha} / 2 \sqrt{2 s^2 / n} && (t_{\alpha} = 2.92, n = 32) \\
 &= 2.92 / 2 \sqrt{2 / 32 (s^2)} \\
 &= .49 \sqrt{s^2} \\
 &= .49 \sqrt{1.49} && (s^2 = 1.49 \text{ for item \#7}) \\
 &= .49 (1.22) \\
 &= .60
 \end{aligned}$$

Next, the means are ranked and differences between adjacent means are computed. (Petersen, 1967)

Iowa	2.78		.09 .22 .00 .88 -- significant
Washington	2.69		
Florida	2.47		
Texas	2.47		
New York	1.59		

The mean for New York responses is significantly larger than the mean for Texas responses (the difference of .88 exceeds the computed L. S. D. value of .60). Significant differences did not occur in the other three comparisons because the differences among the other comparisons of means did not exceed the critical L. S. D. value of .60.