INFORMATION TO USERS

This reproduction was made from a copy of a document sent to us for microfilming. While the most advanced technology has been used to photograph and reproduce this document, the quality of the reproduction is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help clarify markings or notations which may appear on this reproduction.

1. The sign or “target” for pages apparently lacking from the document photographed is “Missing Page(s)”. If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure complete continuity.

2. When an image on the film is obliterated with a round black mark, it is an indication of either blurred copy because of movement during exposure, duplicate copy, or copyrighted materials that should not have been filmed. For blurred pages, a good image of the page can be found in the adjacent frame. If copyrighted materials were deleted, a target note will appear listing the pages in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed, a definite method of “sectioning” the material has been followed. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For illustrations that cannot be satisfactorily reproduced by xerographic means, photographic prints can be purchased at additional cost and inserted into your xerographic copy. These prints are available upon request from the Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases the best available copy has been filmed.
Davey, Kathleen Brady

INSTRUCTIONAL EVALUATION FOR THE PURPOSE OF IMPROVEMENT AND DEVELOPMENT: SELF-REPORTED PRACTICES OF A FACULTY WITHIN A MULTIVERSITY

The Ohio State University Ph.D. 1985

University Microfilms International 300 N. Zeeb Road, Ann Arbor, MI 48106
PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy. Problems encountered with this document have been identified here with a check mark \(\checkmark\).

1. Glossy photographs or pages ______
2. Colored illustrations, paper or print ______
3. Photographs with dark background ______
4. Illustrations are poor copy ______
5. Pages with black marks, not original copy ______
6. Print shows through as there is text on both sides of page ______
7. Indistinct, broken or small print on several pages \(\checkmark\)
8. Print exceeds margin requirements ______
9. Tightly bound copy with print lost in spine ______
10. Computer printout pages with indistinct print ______
11. Page(s) _______ lacking when material received, and not available from school or author.
12. Page(s) _______ seem to be missing in numbering only as text follows.
13. Two pages numbered _______. Text follows.
14. Curling and wrinkled pages ______
15. Dissertation contains pages with print at a slant, filmed as received ______
16. Other__________________________________________

__________________________________________

__________________________________________

University
Microfilms
International
INSTRUCTIONAL EVALUATION FOR THE PURPOSE OF
IMPROVEMENT AND DEVELOPMENT: SELF-REPORTED PRACTICES
OF A FACULTY WITHIN A MULTIVERSITY

DISSertation

Presented in Partial Fulfillment of the Requirements for
the Degree doctor of Philosophy in the Graduate
School of The Ohio State University

By

Kathleen Brady Davey, B.S., M.A.

The Ohio State University

1985

Reading Committee:
Donald P. Sanders
James W. Altschuld
Robert Silverman

Approved By
Advisor

Department of Educational
Policy and Leadership
Vita

June 2, 1944 ...................... Born - Omaha, Nebraska

1974 ............................ B. S., University of Nebraska at Omaha, Omaha, Nebraska

1980 - 1981 ..................... Graduate Assistantship, Meyer Children's Rehabilitation Institute, University of Nebraska Medical Center, Omaha, Nebraska

1981 ............................. M. A., University of Nebraska at Omaha, Omaha, Nebraska

1982 - 1985 ..................... Graduate Research Associate, Instructional Development and Evaluation, The Ohio State University, Columbus, Ohio

FIELDS OF STUDY

Major Field: Educational Policy and Leadership

Studies in: Educational Development
            Educational Research and Evaluation
            Higher Education
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>VITA</td>
<td>ii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vi</td>
</tr>
<tr>
<td>CHAPTER</td>
<td></td>
</tr>
<tr>
<td>1 INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Instructional Decision-Making in Higher Education</td>
<td>1</td>
</tr>
<tr>
<td>Instructional Evaluation Programs and Instructional Decision-making</td>
<td>4</td>
</tr>
<tr>
<td>Problem Background</td>
<td>5</td>
</tr>
<tr>
<td>Statement of the Problem</td>
<td>7</td>
</tr>
<tr>
<td>Research Design</td>
<td>9</td>
</tr>
<tr>
<td>2 LITERATURE REVIEW</td>
<td>11</td>
</tr>
<tr>
<td>Introduction</td>
<td>11</td>
</tr>
<tr>
<td>Organizational and Decision-Making Theory</td>
<td>11</td>
</tr>
<tr>
<td>Descriptive Studies on Faculty Evaluation Practices</td>
<td>18</td>
</tr>
<tr>
<td>Faculty Development Practices</td>
<td>25</td>
</tr>
<tr>
<td>Student Ratings of Teaching</td>
<td>28</td>
</tr>
<tr>
<td>General Instruction Evaluation Guidelines</td>
<td>29</td>
</tr>
<tr>
<td>Summary</td>
<td>31</td>
</tr>
<tr>
<td>3 DESIGN AND METHODOLOGY</td>
<td>34</td>
</tr>
<tr>
<td>Introduction</td>
<td>34</td>
</tr>
<tr>
<td>Sample</td>
<td>35</td>
</tr>
<tr>
<td>Data Collection</td>
<td>38</td>
</tr>
<tr>
<td>Data Analysis and Interpretation</td>
<td>40</td>
</tr>
</tbody>
</table>

iii
4 FINDINGS......................................................................................................... 43
Section A: Context........................................................................................46
Section B: Formal Instructional Evaluation..........................................59
Promotion and Tenure Reviews............................................................... 67
Written Student Evaluations......................................................................77
 Formats in Use.............................................................................................77
 Rating Items...............................................................................................78
 Descriptive Items......................................................................................79
 Reasons for Using Written Student Evaluations................................81
 Perceived Value of Written Student Evaluations for Improvement Purposes........................................................................... 87
 Process of Use...............................................................................................90
 Reported Changes..........................................................................................92
 Factors Related to the Use of Written Student Evaluations
 Faculty Rank......................................................................................  ...94
 Policy..............................................................................................................96
 Annual Performance Reviews.................................................................97
 Analysis of Student Performance on Tests.......................................100
 Departmental Interviews/Surveys of Graduating Seniors............102
 Discussions with Individual Faculty and Teaching
 Assistants for the Purpose of Critique............................................105
 In class Observations by Peers..............................................................107
 Group Discussions with Teaching Faculty for Purpose of Critique........................................................................................................109
 Curriculum Review Committees............................................................111
 Discussions with Individual Students or Groups of Students
 for the Purpose of Critique......................................................................114
 Individually Requested Review of Materials....................................115
 Consultation with a Teaching Improvement Consultant................116
 Faculty Development Committee...........................................................117
 Colleague Rating Form...............................................................................118
 Section C: Informal Instructional Evaluations.....................................119
 Self-Assessment......................................................................................120
 Colleague Assessment................................................................................124
# 5 INTERPRETATIONS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>128</td>
</tr>
<tr>
<td>Faculty as Instructional Decision-Makers</td>
<td>128</td>
</tr>
<tr>
<td>Instructional Evaluation as a Concept</td>
<td>129</td>
</tr>
<tr>
<td>Scenario</td>
<td>131</td>
</tr>
<tr>
<td>Problems with Current Use of Student Ratings</td>
<td>134</td>
</tr>
<tr>
<td>Historical perspective of student ratings</td>
<td>134</td>
</tr>
<tr>
<td>Ratings inadequate for improvement purposes</td>
<td>136</td>
</tr>
<tr>
<td>Ratings inadequate for P&amp; T purposes</td>
<td>137</td>
</tr>
<tr>
<td>Issues Associated with Current Evaluation Practices</td>
<td>143</td>
</tr>
<tr>
<td>Role of Informal Evaluation</td>
<td>144</td>
</tr>
<tr>
<td>An Expanded Notion of Reflection-in Action</td>
<td>147</td>
</tr>
<tr>
<td>Collective Critique</td>
<td>149</td>
</tr>
<tr>
<td>Integrated Critique</td>
<td>151</td>
</tr>
<tr>
<td>Organizational Rewards, Incentives, Norms</td>
<td>152</td>
</tr>
<tr>
<td>Summary</td>
<td>162</td>
</tr>
</tbody>
</table>

## APPENDICES

<table>
<thead>
<tr>
<th>Appendix</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Percentage of Time Reported in Each Activity Area by Each Faculty Participant</td>
<td>173</td>
</tr>
<tr>
<td>B. Faculty Background Sheet</td>
<td>176</td>
</tr>
<tr>
<td>C. Definitions of Codes Used for Organizing Faculty Interview Data</td>
<td>177</td>
</tr>
<tr>
<td>D. Sample Page of Computer Printout for Organized Text Categories</td>
<td>182</td>
</tr>
</tbody>
</table>

BIBLIOGRAPHY: 165
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Percentage Reporting Source Usage</td>
<td>20</td>
</tr>
<tr>
<td>2</td>
<td>Comparison of General Findings from Seven Surveys on Faculty Evaluation</td>
<td>21</td>
</tr>
<tr>
<td>3</td>
<td>Distribution of Faculty Population Across Academic Units and Ranks</td>
<td>37</td>
</tr>
<tr>
<td>4</td>
<td>Distribution of Faculty Sample Across Academic Units and Ranks</td>
<td>37</td>
</tr>
<tr>
<td>5</td>
<td>Average Percentage of Time Reportedly Spent in Each Activity Area According to Activity Profile Groups</td>
<td>48</td>
</tr>
<tr>
<td>6</td>
<td>Average Teaching Load Index (TLI) and Average * of Courses Taught by Profile Group</td>
<td>50</td>
</tr>
<tr>
<td>7</td>
<td>Percentage of Each Faculty Rank in Each Profile Group</td>
<td>52</td>
</tr>
<tr>
<td>8</td>
<td>Formal Instructional Evaluation Processes Reported by Faculty</td>
<td>60</td>
</tr>
<tr>
<td>9</td>
<td>Instructional Evaluation Methods used by Each Faculty Member in the Sample</td>
<td>62</td>
</tr>
<tr>
<td>10</td>
<td>Regularity of Occurrence</td>
<td>64</td>
</tr>
<tr>
<td>11</td>
<td>Types of Instruments and Formats Used for Written Student Evaluations</td>
<td>78</td>
</tr>
<tr>
<td>12</td>
<td>Reasons for Using Written Student Evaluations of Teaching</td>
<td>83</td>
</tr>
<tr>
<td>13</td>
<td>Negative Comments Related to Rating Items and Positive Comments Related to Descriptive Items</td>
<td>88</td>
</tr>
</tbody>
</table>
CHAPTER I: INTRODUCTION

Instructional Decision-Making in Higher Education

Instruction is a major mission of all colleges and universities. In fact, several sources have reported that, regardless of the type of higher education institution, faculty spend the largest part of their time in instruction-related activities (Hildebrand et al., 1971; Baldridge et al., 1978; Ladd & Lipsett, 1978). Even in those institutions classified as "research" or "doctoral-granting" universities, faculty have reported spending, on the average, the largest proportion of their time fulfilling their instructional role (Baldridge et al., 1978; Sell, 1984).

Within institutions of higher education, faculty have been given the primary decision-making authority over the planning, organization, and delivery of instruction.

The faculty is concerned with curriculum, degree requirements, instructional procedure, research activity, faculty status, student academic performance, and the utilization of instructional resources (Millet, 1968, pp. 11-12).

The locus of this instructional decision-making is at the individual department and course level. It is within the context of departments and courses that faculty members decide: 1) what subject matter will be included or excluded in a course or academic program; 2) how subject
matter will be organized within and across courses; 3) how courses will be delivered; and 4) how learning will be evaluated.

These instructional decisions are not of minor significance. The faculty's instructional decisions significantly affect the daily and future lives of university students and the academic reputations of programs, departments, and entire institutions. Further, the subject matter included in and excluded from our university curricula forms the knowledge base and shapes the world view of our college-educated population. Thus, instruction-related decisions made today can affect society-at-large tomorrow.

**Decision-Making As an Activity**

Among decision theorists, decision-making is often defined as a process through which one(s) makes a conscious, rationally based choice(s) from among alternatives which leads to some action(s) (Simon, 1957). Decision theorists also claim that human judgment is inseparable from the decision-making process (MacCrimmon & Taylor, 1977). In the context of decision-making, the act of judging refers to the consideration of some set of facts, evidence, reasons or grounds when selecting a choice for action. It is the judgment-making aspect of decision-making that is most closely related to the concept of "rationally based" choice. As one author points out, choices grounded in reasons or evidence discriminate a judgment from a mere guess, hunch, or conjecture (Greene, 1971).

Further, it has been said that "decision-making, problem-solving, and judgment imply both thought and action (Miller et al., 1960). Indeed they form a bridge between thought and action. Decision-making can be defined
as the process of thought and action that culminates in choice behavior" (MacCrimmon, 1973).

Few, if any, human endeavors occur within some purely rational decision-making context. Rather, actions are taken from relative degrees of conscious, thoughtful, reason-based choice-making. March and Simon (1958) pointed out that:

Most human decision-making, whether individual or organizational, is concerned with the discovery and selection of satisfactory alternatives; only in exceptional cases is it concerned with the discovery and selection of optimal alternatives (p. 140-141).

This "satisficing" phenomenon is due, in part, to the realities of most practical settings. The demands of complex living and working environments do not always allow for identifying, defining, and considering alternative choices to the questions and problems which must be addressed. In addition, the complexity of these problems or questions does not always allow for adequately "knowing" and judging all possible alternative actions. Therefore, humans behave rationally only relative to some set of "given" characteristics of the situation (March and Simon, 1958).

These "givens" include knowledge of assumptions about future events, knowledge of alternatives available for action, knowledge of consequences attached to alternatives--knowledge that may be more or less complete--and rules or principles for ordering consequences or alternatives according to preference (pp. 150-151).

In other words, the more limited the alternatives for action and the more certain the standards for assessing the alternatives, the easier the choices and more likely people are to make them on a rational basis.
Obviously the realities of our higher education’s complex teaching and learning environments, and the instructional problems and situations facing faculty, do not allow for making purely rational instructional choices. Even so, conscious, thoughtful, reason-based instructional choices are a worthy and reasonable goal. As Mann (1973) pointed out:

Everyone agrees that public (and private) decisions should be more "rational," more reasoned, and more reasonable. Rationality is very commonly invoked as a justification for decisions, partly because of our consensus on its desirability, partly because of its symbolic value, and partly because of its heuristic value (p. 21).

**Instructional Evaluation Programs and Instructional Decision-Making**

Within institutions of higher education, numerous formal procedures have been instituted for the purpose of gathering information about instructional practices in order to support and facilitate rationally based decision-making. These formal procedures are traditionally grouped under the general rubric of instructional evaluation (Doyle, 1983). Broadly defined, instructional evaluation can be viewed as means through which information about instruction-related activities is collected, interpreted, and used to make judgments and subsequent decisions about the selection, retention, maintenance of, or changes in instructional programs, personnel, curricula, courses, methods, or specific learning activities.

Within colleges and universities, instructional evaluation occurs most formally and directly through faculty performance reviews and individual course evaluations. Each of these evaluation processes have been instituted, at least in part, for the purpose of serving instructional
decision-making and instructional improvement (Boyd & Schietinger, 1976).
Several surveys have been conducted on administrators' use of various
forms of instructional evaluation information, but little inquiry has been
directed toward understanding if, how, and to what extent faculty acquire
and use information about instruction-related activities for instructional
decision-making.

Problem Background

In the last 20 years, institutional faculty performance evaluation
practices have been the focus of a number of major survey studies (Gustad,
The specific questionnaires varied across these studies. However, in
general, each survey was conducted to determine whether formal
performance reviews occur, how they're conducted, what information about
instructional performance is collected, for what decision-making purposes
instructional information is collected, and how it is weighted in that
decision-making.

All of these studies had similar designs:

- They used mail questionnaires to collect information about
  formal institutional practices.

- They included all types of U.S. institutions of higher education,
  i.e., two-year community colleges, four-year public and private
  liberal arts colleges, public and private doctoral-granting
  institutions.

- They queried primarily administrators, i.e., presidents, central
  administrators, deans, or department chairs.
Similar studies were conducted by Seldin (1975, 1978, 1984), but he surveyed only private and public four-year liberal arts colleges rather than all institution types.

These survey studies span 20 years of institutional practices, and their reported results seem to confirm the following findings:

- Institutions claim that instructional improvement and development is a primary purpose for collecting information about faculty performance. Yet, outside of the promotion, tenure, and retention processes there is little description as to how, if, or by whom the information is actually used for improving and developing instructional practices.

- Faculty evaluation practices particular to doctoral-granting institutions are consistently reported as somewhat different from those operating within other types of higher education institutions.

- There has been a substantial and growing increase in the reported institutional use of systematic student ratings of instructors and courses within all institution types.

- The role that faculty play in reporting information about their own performance has increased over the past decade.

In addition to these descriptive studies of institutional practices, numerous handbooks and manuals have been written on how faculty performance evaluations ought to be conducted (Dressel, 1961; Eble, 1970; Miller, 1972; Cronbach, 1973; Dressel, 1976; Genova, 1976; Centra, 1979; Doyle, 1983; Braskamp et al., 1984). All of these sources include lengthy sections on how to evaluate instructional performance.

In contrast to faculty performance evaluation practices, descriptive studies on course evaluation practices within higher education are almost non-existent. Many articles and books have been written on how to conduct
course evaluations, and detailed methods and instruments have been
developed to assist in course evaluation efforts. Of these methods and
instruments, student evaluation of teaching instruments and in-class
observation tools are among the most frequently researched and cited.
Although there is an exhaustive list of studies conducted for the purpose of
testing the reliability and validity of particular information sources and
course evaluation instruments (Marsh, 1977; Shingles, 1977; Centra, 1977;
Doyle and Crichton, 1978; Uranowitz et al., 1978; Ory et al., 1980; Braskamp
et al., 1981), little research has focused on if and how faculty members
actually use such methods in instructional decision-making.

The literature on formal faculty development programs deals, in part,
with how institutions practice instructional evaluation for the purpose of
Improving instructional practices. Separate survey questionnaires
administered by Centra (1976) and Tannenbaum (1985) included items on
estimated usage and perceived effectiveness of specific Instructional
analysis or assessment practices on efforts to improve instruction. In both
cases, the survey respondents were administrators of faculty development
programs, not faculty users themselves. Further, the questionnaires were
designed to encompass a complete range of faculty development practices
rather than to focus specifically on instructional evaluation and its role in
Instructional decision-making.

Statement of the Problem

Substantial institutional resources are directed toward formal
Instructional evaluation and faculty development programs. A considerable
amount of money and time is devoted to student evaluation of teaching
efforts and performance reviews alone within higher education. Institutions claim that instructional improvement and quality is one of the primary purposes of these formal programs which support instructional evaluation efforts. To date, inquiries have focused primarily on how and to what extent administrators within higher education view and use instructional evaluation information in personnel decision-making rather than on how faculty view and use instructional evaluation information in instructional development and improvement choices.

Personnel decisions can obviously contribute to the institution's instructional quality and improvement by eliminating the "poor" instructors and retaining "good" ones. Nonetheless, it is the faculty who are faced with the myriad of daily, weekly, monthly, quarterly, etc., instructional decisions that have a much more direct impact on instructional quality.

It is the faculty who are primarily "in charge" of instructional improvement and development efforts within colleges and universities. This is particularly true in doctoral-granting universities. Even so, we have little empirical evidence as to whether, how, and to what extent faculty members in multiversities actually acquire, value, and use information about their instruction-related activities in an effort to evaluate and continuously improve their instructional practices. We, therefore, also do not understand if or how well current formal instructional evaluation programs and organizational arrangements in multiversities are actually serving the faculty's instructional decision-making needs and improvement efforts.
Research Design

Because university-level instruction is essentially a faculty controlled activity, it is the faculty who must be queried if one is interested in understanding to what extent and how instruction-related activities are evaluated for the purpose of developing and improving instructional practices. The faculty's perspective of the process of instructional evaluation needs to be understood if organizations are to be able to adequately support and facilitate the achievement of instructional improvement and quality.

This study was designed based on the assumption that the faculty's perspective and description of their actual involvement in instructional evaluation activities could effectively be explored by talking with individual faculty members. It was reasoned that individual interviews would allow faculty to describe in detail their past and current involvement in instructional evaluation and improvement activities and their actual use and value of particular types and sources of information for instructional decision-making. A semi-structured interview strategy was chosen as the primary data-collection methodology.

Research Questions

The general focus of the study was to investigate whether, how, and to what extent faculty acquire, value, and use information about their instruction-related activities in an effort to evaluate and continuously improve their instructional practices. More specific questions guiding the inquiry were these:

1. What do faculty do to evaluate their courses and instructional performance?
2. How and to what extent do faculty acquire information about their instruction-related performance?

3. Of the information available to them, what do faculty value and find useful for improving their instructional practices?

4. What have faculty changed as a result of instructional evaluation information?

5. Are there specific factors that seem to be related to if and how faculty participate in instructional evaluation activities?
CHAPTER 2: LITERATURE REVIEW

Introduction

The literature related to the questions surrounding the faculty's involvement in instructional evaluation and decision-making can be grouped into five main categories:

1. Organizational and decision-making theory particular to higher education;
2. Descriptive studies of how faculty evaluation is practiced within higher education.
3. Articles on faculty development practices;
4. Studies on student evaluation of teaching;
5. General guidelines for conducting instructional evaluation for development and improvement purposes.

Organizational and Decision-Making Theory

The literature on organizations and decision-making provides us with several "structural" models and metaphors for understanding universities as organizations. Each of these, in turn, has implications for related governance and decision-making practices.

Models

Among the most frequently cited "models" for describing the organizational characteristics of higher education institutions are: bureaucratic (Weber, 1947; Stroup, 1966; Blau, 1973); bureaucratic-
professional (Parsons, 1947; Blau and Scott, 1962); collegial (Goodman, 1962; Millett, 1968); and political (Baldridge, 1976; Baldridge et al., 1978). Following is an abbreviated description of some of the characteristics included in each model:

**Bureaucratic**: networks of social groups dedicated to limited goals and organized for maximum efficiency. Authority is derived from appointed formal position. The structure is hierarchical and is connected by formal chains of command and systems of communication (Baldridge et al., 1978).

**Bureaucratic-professional**: authority is derived on the basis of what an actor "knows" and can do, rather than from official position. The workers are deemed "professionals" and therefore control their daily activities without interference or oversight by superiors. There is, however, a surrounding bureaucratic structure which regulates resource allocation, original assignments, scheduling of activities across organizational levels, etc.

**Collegial**: the university is viewed as a collegium or community of scholars. The faculty administers its own affairs with bureaucratic officials having little influence (Baldridge, et al., 1978).

An organization in which functions are differentiated and in which specialization must be brought together, or the coordination, if you will, is achieved not through a structure of superordination and subordination of persons and groups but through a dynamic of consensus (Millett, 1962).

**Political**: "There is a complex social structure that generates multiple pressures, there are many sources and forms of power and pressure that impinge on decision-makers, there is a legislative stage that translates
these pressures into policy, and there is a policy execution phase that generates feedback and potentially new conflicts” (Baldridge et al., 1978).

A slightly different system of decision-making is inferred in each of these organizational models. In particular, each model suggests slightly different realities as to “who makes what decision” and through what process(es). For example, a bureaucratic perspective of the university would suggest that faculty are simply members of an authoritative hierarchy and as such have the authority to make decisions in only very limited and clearly defined areas. It suggests that those who are superior in administrative position monitor and "supervise" the faculty's organizational activities.

On the other hand, a collegial perspective suggests that there is not a decision-making hierarchy and that the faculty as a group, collectively, decide what policy to enact and what actions to take.

Obviously, all of the perspectives offered in these organizational models are to some degree operating in most institutions of higher education. For example, although faculty are members of a bureaucratic hierarchy, they may be given wide latitude over controlling their daily activities. In many cases, they convene as colleagues and collectively decide a variety of matters, but they apparently have little authority over other matters, such as budgets and resource allocations. (Millet, 1968).

While each of these theories of universities as organizations offers us conceptual frameworks for viewing the authority and governance structures of multiversities, none of them have focused attention exclusively on the realm of instructional decision-making. Organizational studies in the area of higher education have focused instead on administrative decision-making,
with the most descriptive detail being offered on the role of university presidents (Millet, 1968; Cohen & March, 1974; Baldridge et al., 1978).

Metaphors

Recent metaphors which have gained prominence in higher education organizational literature are universities as “organized anarchies” (Cohen, March & Olsen, 1972) and universities as “loosely coupled systems” (Weick, 1976).

The view of universities as organized anarchies reflects an attempt to illustrate the unique organizational characteristics of our colleges and universities. Ecker (1980) pointed out that most of the work in organizational theory grew out of the paradigms of administrative science which were more closely associated with business and industry than educational settings.

Most management theory assumes clear goals, technologies that are easily understood, and rather predictable involvement in organizational decision-making. In colleges and universities these conditions do not apply (pp. 24-25).

The imagery of universities as organized anarchies offers an alternative to viewing them within the more traditional organizational constructs. Organized anarchies, according to Cohen et al. (1972), exhibit general properties which they label problematic preferences, unclear technology, and fluid participation.

These authors view universities as organizations which: 1) rather than defining organizational problems and rationally identifying and selecting preferred actions, they wait for members to act and then identify apparent decision preferences from action; 2) “manage to survive and produce,” but
their members often don't understand very well the processes required to perform well in their roles, e.g., faculty often don't have a grounding in designing and delivering instruction; and 3) have participants whose participation in the organization and/or in specific organizational activities is often transitory which produces a high degree of internal organizational uncertainty.

In concert with this non-traditional organizational metaphor, Cohen, March and Olsen offer a parallel model for describing the decision-making process(es) within organized anarchies. They call it a "garbage-can" model of organizational decision-making.

Although it may be convenient to imagine that choice opportunities lead first to the generation of decision alternatives, then to an examination of their consequences, then to an evaluation of those consequences in terms of objectives, and finally to a decision, this type of model is often a poor description of what actually happens. In the garbage can model...a decision is an outcome or interpretation of several relatively independent streams within an organization (e.g., problems, solutions, participants, and choice opportunities). (Cohen, March & Olsen, 1972, p. 3)

Closely related to the ideas associated with the concepts of organized anarchies and a garbage can model of organizational decision-making is the concept of "loosely coupled" systems (March & Olsen, 1975; Meyer and Rowan, 1975; Weick, 1976). These theorists claim that if we look at any educational organization from a system's perspective we find entities that are different in many ways from other formal, work-oriented organizations. In "tightly coupled" systems it is thought that the organization is primarily connected through the technical core of the organization and the authority of office. In educational organizations, and especially in multiversities, it
has been argued that these organizational components function much more independently.

...in the case of technical couplings each element is some kind of technology, task, subtask, role, territory and person, and the couplings are task-induced. In the case of an authority as the coupling mechanism, the elements include positions, offices, responsibilities, opportunities, rewards, and sanctions and it is the couplings among these elements that presumably hold the organization together. A compelling argument can be made that neither of these coupling mechanisms is prominent in educational organizations found in the United States (Weick, 1976, p.4).

There is a great deal of evidence that educational organizations (at least in America) lack internal coordination, especially of the content and methods of what is presumably their main activity—instruction. This main activity tends to be removed from the control of the organizational structure, both in its bureaucratic and collegial aspects. This property of educational organizations, among others, has led March to apply the term loosely coupled to educational organizations. By this term he means that structure is disconnected from activity, and activity is disconnected from its effects (Meyer and Rowan, 1975, p. 1).

The concept of loosely coupled systems has particularly important implications when considering instructional decision-making, because it supports the premise that one must focus on individual faculty members to understand how and to what extent instructional evaluation that is aimed at improvement and development operates within multiversities.

The separate works of Simon (1957), Millett (1968) and Sell (1980) offer further theoretical support for the need to focus on faculty members when studying instructional evaluation and decision-making aimed at instructional improvement and development. Simon refers to the persons
who physically carry out an organization's objectives as "operative employees." In colleges and universities it is clearly the faculty who enact the organization's instructional objectives. In fact, it is the faculty who define, enact, and in large part evaluate the achievement of instructional objectives within higher education. Simon (1957) says:

In the study of organizations, the operative employee must be at the focus of attention, for the success of the structure will be judged by his performance within it. Insight into the structure and function of an organization can but be gained by analyzing the manner in which the decisions and behavior of such employees are influenced within and by the organization (p.3).

Millett distinguishes between the role and responsibilities of administrators in higher education compared to the faculty. His distinctions help demonstrate why focusing on administrative behavior within these institutions inadequately informs us regarding instructional decision-making.

[Administrators are concerned with] financial resources to collect and augment, budgets to adopt and execute, accounting records to maintain, supplies to be purchased and stored, buildings to plan and maintain, non-academic employees to be hired and retained, facilities to be utilized and controlled, housing to be provided students, and services to be operated (such as water, heat, light, sewerage, and communication). This enumeration says nothing about the direction of the educational operation itself, including instruction, research, and public service (p. 9).

Finally, Sell (1980) in a review of the organizational literature applied to higher education through 1980, concluded that:
With regard to substance, many studies focus primarily on organizational structure in the formal sense; a number of others deal with goals/outcomes, decision-making and communication processes, organizational personnel (e.g., administrators, faculty, students). Few organizational studies in higher education focus on instructional technology or teaching-learning activities, essential areas of understanding for the administration of colleges and universities (p.37).

Descriptive Studies on Faculty Evaluation Practices

In higher education, instructional evaluation practices have been examined primarily from the perspective of how administrators conduct faculty performance evaluations. In the last two decades, at least seven major surveys of faculty performance evaluation practices have been conducted (Gustad, 1961; Astin & Lee, 1967; Seldin, 1975, 1978, 1984; Boyd & Schietinger, 1976; and Centra, 1977). One was a regional study (Boyd & Schietinger, 1976, Southern Regional Education Board), and the other six drew national samples. Most of these surveys collected data on the evaluation of the faculty members’ total performance, including instruction, research, and public service. A major purpose of each of these inquiries was to describe to what extent and how instructional performance is evaluated within colleges and universities.

Each of these studies defined in slightly different ways the classes of institutions represented in their samples and also used slightly different questionnaire items. In most cases it is therefore not possible to make direct comparisons of the findings across studies. Seldin’s three studies are the one exception. He used a consistent definition of private four-year liberal arts colleges in at least a part of his sample in all three of his surveys and was therefore able to fairly confidently define trends over time
in those colleges. Based on the responses of academic deans, Seldin was able to identify the following trends in faculty performance evaluation practices in private four-year liberal arts colleges over the last ten years.

1. These institutions have increased their emphasis on the performance areas of research, publications, and activity in professional societies.

2. They have decreased their emphasis on length of service in rank, competing job offers and personal attributes.

3. They have increased their use of formal student ratings, classroom visits and self-evaluation as sources of evaluative information. The use of formal student ratings and self-evaluation represents the most significant change in practices.

Table 1 represents the percentage of academic deans in each of Seldin's surveys who reported using each of these three sources of information.
## TABLE 1

Percentage Reporting Source Usage

<table>
<thead>
<tr>
<th>Source</th>
<th>% 1973</th>
<th>% 1978</th>
<th>% 1983</th>
</tr>
</thead>
<tbody>
<tr>
<td>Formal Student Ratings</td>
<td>29.1</td>
<td>53.1</td>
<td>66.6</td>
</tr>
<tr>
<td>Classroom Visits</td>
<td>4.9</td>
<td>11.1</td>
<td>17.3</td>
</tr>
<tr>
<td>Self-evaluations</td>
<td>19.8</td>
<td>36.3</td>
<td>42.5</td>
</tr>
</tbody>
</table>

By comparing the results of his three studies, Seldin (1984) concludes that:

The trend over the past five years is evident. Information gathering is now more structured and systematic. A growing number of institutions are making concerted efforts to formalize and shore-up their evaluative methods (p.44). There is no doubt that private liberal arts colleges in the past decade have been busy dismantling and reconstructing some of their procedures in the faculty evaluation system (p.50).

Table 2 was constructed in order to compare general findings from the seven major surveys on faculty performance evaluation. Even though the differences in the designs of these studies do not allow for technically perfect comparisons, it is possible to at least form general impressions about trends and possible changes over time by comparing their general findings.
### TABLE 2
Comparison of General Findings From Seven Surveys on Faculty Evaluation

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>5. Research</td>
<td>5. Research</td>
<td>5. Supervision of graduate study</td>
<td>5. Personal attributes</td>
<td>5. Research</td>
<td>5. Research and/or creative activity (independent of publications)</td>
<td>5. Research</td>
<td>5. Research</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>5. Supervision of graduate study</td>
<td>5. Supervision of graduate study</td>
<td>5. Activity in professional societies</td>
<td>5. Civic activities</td>
<td>5. Supervision of honors study</td>
<td>5. Supervision of honors program</td>
<td>5. Supervision of honors program</td>
<td>5. Supervision of honors program</td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>--------------</td>
<td>------------------</td>
<td>--------------</td>
<td>--------------------------</td>
<td>--------------</td>
<td>--------------</td>
<td>--------------</td>
</tr>
<tr>
<td>5 most utilized</td>
<td>1. Informal student opinion</td>
<td>1. Chair evaluation</td>
<td>1. Dean evaluation</td>
<td>1. Academic Dean or Vice Pres.</td>
<td>1. Chairman evaluation</td>
<td>1. Chair evaluation</td>
<td>1. Chair evaluation</td>
</tr>
</tbody>
</table>

### TABLE 2 (Continued)
The most obvious trends across these studies are:

1. The administrators' reporting of classroom teaching activities as a major or critical factor in faculty performance evaluations;
2. The fairly consistent inclusion of research, personal attributes, and student advising as a major performance factor;
3. The relatively consistent designation of consultation activities, competing job offers, and supervision of honors programs as unimportant factors;
4. The consistent designation of the chair's evaluation, dean's evaluation, and systematic student ratings as sources of information used in faculty evaluations.

In those studies in which the sample included institutions defined as doctoral-granting, multiversities, or large complex universities (Gustad, 1961; Astin & Lee, 1966; Boyd & Schietinger, 1976; Centra, 1977), the researchers consistently described these institutions as reporting practices somewhat different from the other classes of institutions. This difference was particularly represented in administrators' designation of research and publications as a performance evaluation factor often equal in importance with teaching activities. Administrators in these institutions also tended to report relying, more than other types of institutions, on systematic student ratings as an information source, and less heavily on personal attributes.

There was only one faculty evaluation study in which faculty were included as targeted respondents. Thorne et al. (1976) studied the faculty evaluation practices operating within the Oregon State System of Higher Education (OSSHE). As a part of their study, they surveyed faculty on what
performance areas and information sources they perceived as being influential in faculty performance reviews. Four hundred and eighty-five faculty responded to the survey. Several important findings were reported from this study.

According to the OSSHE study, several of the information sources that are often recommended to be used for the purpose of improving and developing instruction are viewed by faculty as the least influential and least used in administrative performance reviews. For example, the faculty reported perceiving formal appraisals of teaching, syllabi, classroom examinations, service activities, or research activities as among the sources of information least influential in performance evaluations.

We found that the faculty saw the least influence being exerted by formal methods of appraisal, whatever the professional function (p. 26).

The findings from the Oregon Study, when compared to the national and regional surveys, also suggests that what administrators report as the most important performance areas and influential information does not necessarily match the faculty's perceptions. Thorne et al. reported that faculty perceive publications, not measures of teaching performance, as the most influential factor in performance appraisals.

The Oregon researchers also reported a fair amount of variance in the way faculty within a given department perceived faculty evaluation practices. The researchers concluded that the variance represented some aspect of uncertainty which could be derived from:

1) Lack of awareness of procedures or criteria by individual faculty members;
2) Department-specific conditions, i.e., there may be specific departments in which clear communications about indicators have not occurred;

3) Nebulous campus procedures, guidelines and policy statements (pp. 21-22).

As a group, the surveys on faculty performance evaluation practices provide a "macro" view of what information about faculty's instructional performance is collected and claimed by administrators to be used in faculty performance reviews. They therefore provide some basis for inquiring into whether the instructional evaluation information collected and used in faculty performance reviews is the same information that is perceived by faculty as useful in instructional development and improvement efforts. There is some doubt, however, as to whether the information collected and valued for performance reviews is currently serving, or can be expected to serve, both functions.

From an institutional point of view, the emphasis on evaluation for development and the emphasis on performance evaluation are opposite sides of the same coin and may draw upon the same kind of information. From the perspective of the individual faculty member, however, they contrast sharply--faculty development generally connotes enlargement of personal opportunity, while performance evaluation may pose a threat (Boyd & Schleitinger, 1976, p.2).

Faculty Development Practices

Formal faculty development programs are one organizational link between instructional evaluation and instructional improvement efforts. As Goldschmid (1978) explained:
A number of universities have established staff development programs often organized by special units, which have become known under a variety of names, such as higher education advisory and research centers, pedagogical service units and staff or faculty development institutes. By offering a documentation service, courses, workshops, seminars and consultation on teaching, as well as engaging in research on instruction, these units are supposed to provide the impetus of teaching improvement efforts (Goldschmid, 1978, p.233).

A variety of instructional evaluation methods have been used with faculty as a part of formal faculty development activities. Some of these include systematic ratings of instruction, formal assessments by colleagues, classroom visitation by an instructional resource person, and professional and personal development plans for individual faculty members (Centra, 1976).

Faculty development programs have been the focus of "macro" survey studies (Miller & Wilson, 1963; Many, Ellis & Abrams, 1969; Eble, 1970; and Centra, 1976) similar to those which dealt with institutional faculty evaluation practices. In these studies, administrators were surveyed for the purpose of finding out if such programs existed and, if so, what development activities were offered.

The most comprehensive of the faculty performance studies was conducted by Centra (1976). Seven hundred and thirty-four faculty development administrators from doctoral-granting universities, four-year colleges, and two-year colleges responded to a mail questionnaire. They were asked to identify the types of faculty development activities operating
on their campuses and to estimate the proportion of faculty involved in each. They were further asked to rate how effective they thought each activity was. The activities covered in the questionnaire included a full range of development activities such as seminars, sabbaticals, and grants for materials development. It also included a ten-item section for reporting techniques which might have been used to assess, evaluate and provide feedback to faculty about instruction. Centra reports that, in general, respondents from the two-year colleges were more likely than those from four-year colleges and universities to rate instructional analysis and assessment techniques as effective for producing instructional improvement. Since directors of faculty development programs rather than faculty were the informants in this study, the data may or may not accurately reflect actual faculty use and the faculty's perceptions of effectiveness.

Goldschmid's reports of her experiences with faculty development programs cast serious doubts as to whether data reported by faculty development programs can accurately reflect actual practices of the majority of faculty. She claims that the number of faculty who use the services of development centers represents approximately 1% to 10% of the total number of instructors. Furthermore, both Goldschmid and Centra note that the majority of the teachers who use the services of faculty development programs represent the best and most concerned faculty.

Finally, faculty were surveyed directly in one phase of a study dealing with the relationship between institutional policy on faculty evaluation and the degree of faculty involvement in faculty development activities (O'Connell, 1983). Usable data were collected from 95 faculty in four
private liberal arts colleges. The faculty were asked to self-report their involvement in a variety of faculty development activities and the subsequent changes in their teacher behavior. The list of faculty activities specified in the survey included only one obvious instructional evaluation activity, i.e., consultation with a colleague. The other activities specified in the questionnaire were short seminars or workshops, two-three day retreats, attendance at professional meetings, activity toward a degree, etc.

Student Ratings of Teaching

Student ratings of teaching are so universally used on college and university campuses that they seem to merit separate discussion as an instructional evaluation method related to instructional improvement.

In the late 1960s there was an upsurge in the demand from university students for the opportunity to evaluate their professors' performance. This demand led to an exponential rise in the number of student evaluation instruments developed and subsequent studies testing their validity and reliability (McKeachie, Lin & Mann, 1971; Kulik & McKeachie, 1975; Centra, 1977; Feldman, 1977; Marsh, 1977; Whitely & Doyle, 1979; Cohen, 1981). Rotem and Glasman (1979) in a controversial article concluded, after reviewing a large body of research on feedback, that there is a "minimal effect at best of feedback on instructional improvement at the university level (p. 497)." Other research somewhat contradicts this observation. Two conditions have been identified which contribute to the possibility that feedback from student ratings might lead to changes and improvement. McKeachie & Lin (1971,1975), Jocoby (1977), and Aleamoni (1978) all conducted studies which resulted in their reporting that ratings, when
accompanied by personalized consultation, contribute to instructional improvement. In addition, Centra (1973), McKeachie and Lin (1975) and Payne and Hobbs (1979) conducted studies which led them to conclude that:

For instructors whose self-evaluations were considerably better than were their student ratings, changes in instruction (as assessed by repeated student evaluations) occur after only a half semester. If, in other words, teachers are "unrealistic" relative to their students' view, that is--then they tend to make changes in their instructional practices (Centra, 1979, p. 39).

The literature on student evaluations provides a rich basis for making judgments about the technical quality of particular instruments and the conditions under which student evaluations and specific items are technically reliable as information sources. The research up to now, however, says little about how, under natural conditions, student evaluations are actually used by faculty and whether they are perceived as useful for addressing faculty's instructional improvement concerns.

General Instructional Evaluation Guidelines

General guidelines or strategies for conducting instructional evaluations for the purpose of development and improvement have been outlined in a number of evaluation handbooks (Dressel, 1961; Eble, 1970; Miller, 1972; Cronbach, 1973; Dressel, 1976; Genova, 1976; Centra, 1979; Doyle, 1983; Braskamp et al., 1984). Most of these suggested approaches rest on the assumption that:

In reality, there is but one significant reason for evaluating instruction--to improve the quality of learning and increase the percentage of students who attain the important and agreed upon goals of learning. All else flows out of and is secondary to that central goal (Dressel, 1976, p. 338).
Instructional evaluation for improvement is sometimes referred to as "diagnostic" or formative evaluation. It is usually viewed as any process through which some aspect(s) of instruction is (are) considered for the purpose of ascertaining if something needs to be or could be changed that would lead to improved instructional process(es) or outcome(s).

A great deal of the early work in instructional evaluation methods tended to concentrate on in-class teacher performance and resulted in a large number of teacher behavior observation instruments. Most current leaders in the field now recognize that:

The instructional role is much broader, including all tasks related to teaching. These tasks extend to selection of text, additional reading, preparation and grading of tests, interactions with other faculty members teaching the same course or with those in other departments whose majors take the course. Like the tip of an iceberg, classroom teaching is the most visible part of instruction, and concentration solely on that visible portion can result in tragedy (Dressel, 1976, p. 337-338).

It is usually recognized that focusing only on in-class teacher behaviors is a necessary but insufficient means of practicing instructional evaluation for improvement purposes. Braskamp et al. describe some commonly accepted guidelines for conducting instructional evaluation programs aimed at improvement and development in their Guidebook for Evaluating Teaching, Part I, reprinted in the NACTA Journal, 1983-1984. These guidelines are listed below, including questions which the authors propose should be asked in order to judge whether such evaluations are operating.
1. Information collected for improvement is collected for the instructor only. **Question:** Do faculty have the freedom to collect evaluative information for their private use?

2. Information is frequently and informally collected. **Question:** Are faculty collecting enough information to monitor their progress?

3. Evaluation tied to self-development maximizes its long range utility. **Question:** Does the faculty member accept the principle that self-evaluation of teaching is a necessary condition for change and growth?

4. Information collected that is highly detailed, diagnostic, and focused on specific teaching behaviors and course characteristics (e.g., test, text) increases the usefulness of the information. **Question:** From the information collected, does the instructor know specific strengths and weaknesses?

5. Information shared with another often increases the usefulness of the information. **Question:** Do faculty who desire to discuss their teaching have opportunities to receive consultative assistance?

The perspective offered in the instructional evaluation handbooks offer us a basis for operationalizing instructional evaluation for improvement purposes and a conceptual framework for ascertaining whether such processes are occurring in a particular setting.

**Summary**

The traditional organizational literature offers a variety of ways to view the organizational structure of colleges and universities. The more recent metaphors of universities as "organized anarchies" and "loosely coupled
systems focus on the unique organizational characteristics of complex universities. No matter what lens is used for viewing universities as organizations, it becomes clear that the locus of control for instructional decision-making rests with the faculty. As such, it is the faculty who have the onus and challenge of evaluating instruction for development and improvement purposes.

Instructional evaluation as a process has been studied primarily from the perspective of administrators through descriptive surveys on institutional faculty evaluation and faculty development programs. While these surveys have provided a perspective of 1) what information is claimed to be collected and used for personnel decisions, and 2) what instructional evaluation methods are used as a part of formal faculty development programs, they provide no reports of how faculty in natural settings view and use instructional evaluation as a means for improving and developing their instructional practices.

Finally, the faculty performance evaluation studies indicate that student evaluation of teaching is viewed as a predominant and growing instructional evaluation tool as practiced within higher education. Further, handbooks on instructional evaluation offer useful guidelines for using student evaluation and other evaluation methods when the aim is instructional improvement. Again, however, the literature does not describe how any of these methods are viewed by faculty in natural settings, or if and to what extent they use them for instructional improvement purposes.
PLEASE NOTE:

This page not included with original material. Filmed as received.

University Microfilms International
CHAPTER 3: DESIGN AND METHODOLOGY

Introduction

This study was designed based on the following assumptions:

- If instructional evaluation aimed at instructional improvement is occurring, the faculty are primary actors in the process.
- While institutions of higher education share common organizational characteristics, the instructional role of faculty as enacted in multiversities, or "doctoral-granting" universities, is somewhat different from that in other institution-types because of the organization's shared emphasis on research as a mission.
- The literature on institutional faculty evaluation practices supports the premise that, as an institution class, doctoral-granting institutions merit separate study.
- So little is currently known about how faculty view and use instructional evaluation aimed at improvement and development that a research methodology is called for which allows a descriptive base to be built. A reasonable juncture for building this description is with the self-reported viewpoints and practices of faculty members.
A case study approach, using semi-structured interviews, was selected as a research methodology. The purpose of the interviews was to allow a sample of faculty within a doctoral-granting university to describe in detail if and how they had been involved in collecting and using information about their instructional practices for the purpose of instructional development and improvement. The intent of using semi-structured interviews as a means of data collection was to avoid unduly "shaping" the responses of the faculty. The strategy was chosen to facilitate the faculty's description of actual practices and to permit probing for details on how information was actually used and valued.

The size and organizational complexities of doctoral-granting universities make them problematic settings for selecting a faculty interview sample that might provide both depth and breadth to the resulting findings. It was decided to sacrifice representativeness for the opportunity to provide richer detail and more trustworthy findings about faculty within one setting. Therefore, one college within a doctoral-granting institution was chosen as the site for this study. This choice allowed the researcher to build into the design several of the methodological considerations that Guba (1980) has outlined for increasing the trustworthiness of qualitative research, i.e., peer debriefing, collection of referential adequacy materials, establishment of audit trails, member checks.

Sample. One college within the Ohio State University was selected as the site for the study. The college was considered to be representative of colleges within doctoral-granting universities. Its faculty represent a range of academic disciplines and ranks and are involved in instruction, research and service activities. The college provides both graduate and
undergraduate instruction through a variety of academic programs and instructional settings, e.g., classrooms, seminars, labs and field experiences.

A sample of one-third of the faculty was derived by selecting every third name from an alphabetized list of the college's roster of fulltime faculty teaching courses during Spring Quarter, 1984. The distribution of the 41 selected faculty, across the 11 academic units and three faculty ranks, approximated the distribution of the college faculty population. Table 3 shows, according to academic unit and faculty rank, the distribution of the faculty population in the college. Table 4 shows the distribution of the sample according to the same characteristics.
### TABLE 3
Distribution of Faculty Population across Academic Units and Ranks

<table>
<thead>
<tr>
<th>Academic Unit</th>
<th>Full-Time Faculty</th>
<th>Faculty Rank (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>01</td>
<td>24</td>
<td>19</td>
</tr>
<tr>
<td>02</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>03</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td>04</td>
<td>16</td>
<td>13</td>
</tr>
<tr>
<td>05</td>
<td>13</td>
<td>10</td>
</tr>
<tr>
<td>06</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>07</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>08</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td>09</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>10</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>11</td>
<td>15</td>
<td>12</td>
</tr>
<tr>
<td>Total</td>
<td>125</td>
<td>101</td>
</tr>
</tbody>
</table>

*Rounded to nearest whole percent

### TABLE 4
Distribution of Faculty Sample across Academic Units and Ranks

<table>
<thead>
<tr>
<th>Academic Unit</th>
<th>Full-Time Faculty</th>
<th>Faculty Rank (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>01</td>
<td>8</td>
<td>20</td>
</tr>
<tr>
<td>02</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>03</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>04</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>05</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>06</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>07</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>08</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>09</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>10</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>11</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>Total</td>
<td>41</td>
<td>100</td>
</tr>
</tbody>
</table>

*Rounded to nearest whole percent
Data Collection

The 41 faculty were each invited to participate in one-hour interviews. Only one person declined, and he was replaced by another randomly selected faculty member of the same rank. The faculty were contacted by phone to schedule interview appointments, and they were sent a background information sheet to complete prior to their individually scheduled meetings (See Appendix A).

Two interviewers were present at 38 of the 41 interviews. With the permission of each faculty participant, all interviews were audio-recorded. This approach freed the interviewers to concentrate on the interview process.

The interviews proceeded from the general to the specific. First, the background information sheet was reviewed and discussed to assure that the faculty member's overall teaching responsibilities and assignments were accurately recorded and understood by the interviewers. The faculty were then asked to describe what, if anything, they or others did to obtain information about their instruction-related activities. Each participant was asked probing questions regarding the actual sources, uses, and users of the information. Each was further asked to describe information he or she found useful for making changes and improvements and whether they had ever changed their teaching, courses or curricula as a result of information they or others had acquired. If the faculty member did not mention any evaluation activities with students, colleagues, department chairs or other individuals, the interviewers asked about these parties as possible sources of information, as well as their actual involvement with the interviewee.
If the faculty member reported using instruments or forms for collecting information, the interviewers requested to see a copy and asked when, how often, and for what purposes the instrument(s) was (were) used. The instruments were collected for further analysis later.

The final part of each interview was arranged to give the faculty member the opportunity to reflect on and discuss problems, issues, and recommendations related to teaching, instructional evaluation practices, and the improvement of instruction in the university.

A typical interview included the following line of inquiry:

--Tell us a little bit about the courses you teach. Are they primarily graduate or undergraduate? What does the enrollment tend to be? How are they structured?

--Could you talk a little bit about how you define the areas of instruction, research, service and student advising for yourself and how you distribute your time across those activity areas?

--What, if anything, have you ever done to evaluate your instructional practices? To assess how they're going and what might need to be changed?

--Do you use students to evaluate your instructional practices?

--Have you ever consulted with other faculty about your courses?

--Has your chair been involved in evaluating your courses?

--Is there anything else, formal or informal, that you've ever done to assess your instructional practices?

--Are there any other instructional evaluation activities that go on in your department?

Any time an interviewee mentioned an instructional evaluation activity, he or she was then asked more probing questions such as: What kind of
information was collected through this means? Who was involved in the process? What was done with the information? Did you make any changes as a result?

Data Analysis and Interpretation

The two interviewers "debriefed" one another by discussing the interviews with one another after each interview session. Through this process they attempted to, throughout the course of the study, clarify their perceptions and interpretations of the faculty's comments.

The interviewers collected adequacy referential materials in the form of course descriptions, syllabi, evaluation instruments, and program descriptions. They also consulted departmental registration records in order to verify self-reported course assignment data.

Transcriptions of the interviews were prepared from the audiotapes. A constant-comparative analysis method (Glaser, 1969) was used while reviewing the transcriptions. That is, the transcriptions were reviewed for the purpose of identifying all possible common and divergent themes, constructs, or content categories. These categories were derived from the content of the interviews and were not pre-defined. One of the interviewers reviewed a single transcript and noted all possible content category labels. Each comment was "constantly compared" with the previous comments in order to determine if it was categorically the same or different. The interviewer proceeded with this process for several transcripts, each time trying to "bracket" the conceptual framework derived from the previous transcripts. This process was followed until no new categories seemed to emerge (approximately eight transcripts). The second interviewer reviewed this work, and the two interviewers discussed the categories and confirmed
that all of the important content from the transcriptions was accurately accounted for in one or more of the general construct labels listed and defined in Appendix B.

These derived category labels were not considered sacred, final or mutually exclusive. They were rather viewed as an aid for organizing this massive volume of text into a more orderly, accessible, and useful form than 41 intact interview transcriptions. The original transcriptions were also available for reinspection as needed.

The categories were assigned a symbolic code and all 41 transcriptions were reviewed and manually coded using this scheme. Each coded transcript was then entered into a mainframe computer. Using the WYLBUR program in conjunction with the SCRIPT word processing program, the researchers were able to create two large files and to assign a detailed code to each line of text. The code identified each line by the content category, interviewee, department, rank, tenure status, and a teaching load index. A "list" command then allowed the researchers to merge all the comments related to a particular category and print them by category. This coding and organization allowed the researchers to easily review comments across the sample by category, or to review comments made by a single participant across categories. It also allowed them to quickly access comments by faculty member, department or rank. A sample page of the 22,000-line printout used for analysis is shown in Appendix C.

The two interviewers discussed perceived themes, patterns, and relationships and reached agreements regarding a more refined conceptual framework for reporting the findings as well as reasonable interpretations of the data. An audit trail was established by carefully filing all notes,
transcripts, coding schemes, audiotapes, referential documents, and drafts of all analysis worksheets and reports. The findings and interpretations resulting from this study are reported in the next two chapters.

The purpose of this study was not to gather quantitative measures of faculty practices which would allow for statistical analyses of the "data". The intent, rather, was to understand in as much detail as possible the perceptions and experiences with instructional evaluation of a group of faculty representative of one setting. The chosen methodology, therefore, limits the ability to automatically generalize the findings to other settings. It is believed, however, that the rigor with which the interviews were documented and analyzed, and the detail with which the findings are reported, presents an opportunity for individual readers to confidently judge the applicability of the findings to their own setting. While most statistical analysis methods were not appropriate for interpreting these "data", counting of responses by analytical categories is frequently reported as simply a descriptive characteristic of the group's practices.
CHAPTER 4: FINDINGS

The goal of this chapter is to report, without evaluating, the faculty comments from the interviews which pertain most directly to the research questions associated with the study, i.e.,

1. What, if anything, do faculty do to evaluate their courses and instructional performance?
2. How, and to what extent do faculty acquire information about their instruction-related performance?
3. Of the information available to them, what do faculty value and find useful for improving their instructional practices?
4. What, if anything, do faculty report having changed as a result of instructional evaluation information?
5. Are there specific factors that seem to be related to if and how faculty participate in instructional evaluation activities?

In Chapter 5 some possible interpretations of the findings are proposed, and this chapter (Chapter 4) is intended to provide the reader with the factual basis from which the interpretations were drawn.

Chapter 4 is organized into three main sections: Section A: Context; Section B: Formal Instructional Evaluation; and Section C: Informal Instructional Evaluation. In Section A some of the contextual information which the faculty participants reported on their Background Information Sheet (Appendix A), and which they also discussed in their interviews, is summarized. Sections B and C report the faculty's
descriptions of instructional evaluation as it actually operates in their settings. Although instructional evaluation is usually treated in the literature and discussed among educators as if it were some very technical and complex process, requiring particular expertise, skills, instruments, etc., according to these interviews, instructional evaluation for improvement purposes is not always viewed by faculty as such a formal, purposeful activity. In fact, the faculty participants repeatedly named the acquisition of information through the normal course of teaching, talking, etc., as a means of "instructional evaluation." Therefore, Section B reports the faculty's involvement in and perceptions of formal instructional evaluation activities, and Section C, their perceptions of instructional evaluation as a more natural process.

Direct quotes from the interviews are reported throughout the Findings chapter. The person responsible for each statement in cited using a faculty code number following each quotation.

The instructional evaluation practices as reported in this chapter may not represent a complete description of every instructional evaluation activity in which each faculty member had ever participated. Individual faculty may have failed to recall or to mention every instance that ever occurred over time. Still in all, the practices as described do represent a complete set of activities that the faculty considered important enough to mention. Because probing questions were asked of each interviewee, it is assumed, with a reasonable degree of confidence, that across all faculty participants the major means of instructional evaluation activities as practiced by these faculty were identified. It is also assumed that these
descriptions represent a reasonable approximation of their actual frequency and conditions of use.
Section A: Context

The 41 faculty interviewed in this study represented 11 different academic units within one college at The Ohio State University. Ohio State has one of the largest enrollments of all multiversities in the United States (58,000 plus students) and is organized into one main campus at Columbus, four regional campuses, one Agricultural Technical Institute, 18 colleges, 7 schools, 51 divisions, and 125 academic departments (The Ohio State University Faculty and Staff Directory, 1985).

The 41 faculty members’ combined years of college teaching equalled 586 years, ranging from 1.8 to 38 years, and all three faculty ranks—assistant, associate, and full—were represented in the sample.

The faculty distribute their time across five organizational activity areas: instruction, research, public service, administration, and advising. Based on self-reports of the percentage of time spent in each activity area over an academic year, the interviewed faculty were grouped according to “activity profiles.” Five profiles were identified, with 85% of the sample fitting one of three profile groups, i.e., instruction primary activity (37%), research primary activity (22%), or balance between instruction and research (27%). The profile groups were defined as follows:

Group 1 (primary activity instruction): reported spending at least 15% more time in instruction than in any other single activity area.

Group 2 (primary activity research): reported spending at least 15% more time in research than in any other single activity area.
Group 3 (balanced between instruction and research): reported no more than a 10% difference between time spent in instruction and time spent in research and reported spending at least 15% more time in each of the activity areas of instruction and research than any other single activity area.

Group 4 (primary activity public service): reported spending at least 15% more time in public service than any other single activity area.

Group 5 (balanced between at least three activity areas): reported no more than a 15% difference among time spent in at least three different activity areas.

Table 5 displays by profile group the data regarding the average percentage of time reportedly spent in each activity area. Appendix D displays the individual data from which the group data were derived.
## TABLE 5

Average Percentage of Time Reportedly Spent in Each Activity Area According to Activity Profile Groups

<table>
<thead>
<tr>
<th>Profile Group</th>
<th>n</th>
<th>% of Sample</th>
<th>I</th>
<th>R</th>
<th>PS</th>
<th>ADM</th>
<th>ADV *</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1 (Instruction)</td>
<td>15</td>
<td>37 60 14 8 9 10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group 2 (Research)</td>
<td>9</td>
<td>22 24 51 8 6 11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group 3 (Instruction/Research)</td>
<td>11</td>
<td>27 41 39 5 7 8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group 4 (Public Service)</td>
<td>4</td>
<td>10 14 12 67 3 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group 5 (3-way Balance)</td>
<td>2</td>
<td>5 38 10 20 13 20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* I  =  Instruction  
  R  =  Research  
  PS  =  Public Service  
  Adm  =  Administration  
  Adv  =  Student =Advising
These self-reported data were verified by reviewing college records regarding the number and level of courses taught by each of the faculty participants in academic year 1983-84. Class enrollments and credit hours were also recorded, and a teaching load index for each participant was derived according to the following formula:

\[
\text{Student credit hours} = \text{credit hours designated for a course(s)} \times \text{course enrollment(s)}.
\]

1283 = maximum number of undergraduate student credit hours generated by any single faculty member in the sample.

777 = maximum number of graduate student credit hours generated by any single faculty member in the sample.

12 = maximum number of courses taught by any single faculty member in the sample within academic year 1983-84.

7 = maximum number of new course starts for any single faculty member in the sample. A new course start is defined as a course with a course number discrete from any other course taught by the same faculty member.

\[
\text{Teaching Load Index} = \left( \frac{\text{Undergraduate student credit hours}}{1283} + \frac{\text{Graduate student credit hours}}{777} \times 100 \right) + \left( \frac{\text{Number of courses taught}}{12} + \frac{\text{number of new course starts}}{7} \times 100 \right)
\]

The average teaching load index and average number of courses taught by each profile group appears in Table 6.
TABLE 6

Average Teaching Load Index (TLI) and Average # of Courses Taught by Profile Group

<table>
<thead>
<tr>
<th>Group</th>
<th># of Courses</th>
<th>TLI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>5</td>
<td>96</td>
</tr>
<tr>
<td>Group 2</td>
<td>2</td>
<td>51</td>
</tr>
<tr>
<td>Group 3</td>
<td>3</td>
<td>61</td>
</tr>
<tr>
<td>Group 4</td>
<td>1</td>
<td>17</td>
</tr>
<tr>
<td>Group 5</td>
<td>2</td>
<td>44</td>
</tr>
</tbody>
</table>
As a group, those persons reporting instruction as their primary activity area represented the highest average teaching load index. Those reporting a balance between instruction and research represented a lower TLI than the instruction group, but higher than the research group, etc. There were a few cases where faculty reported either instruction or research as a major activity area but whose TLI reflected the reverse situation. It was assumed that in these cases the self-reported data most likely did not reflect actual circumstances, or the individuals had an atypical way of performing the various activities. Because data were not collected to assess the relative productivity (outcomes) of faculty time allocated to research, service or administration, these comparisons with teaching load could not be made.

In Table 7 the number of faculty in each profile group are listed by rank.
TABLE 7

Percentage of Each Faculty Rank in Each Profile Group
(Faculty Rank = Full Professor, Associate Professor, or Assistant Professor)

<table>
<thead>
<tr>
<th>Group</th>
<th>Faculty Rank</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Full % of n</td>
<td>Associate % of n</td>
<td>Assistant % of n</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rank</td>
<td>Rank</td>
<td>Rank</td>
<td>Rank</td>
</tr>
<tr>
<td>Group 1</td>
<td>8 44</td>
<td>2 14</td>
<td>5 56</td>
<td></td>
</tr>
<tr>
<td>Group 2</td>
<td>4 22</td>
<td>1 7</td>
<td>4 44</td>
<td></td>
</tr>
<tr>
<td>Group 3</td>
<td>3 17</td>
<td>8 57</td>
<td>0 00</td>
<td></td>
</tr>
<tr>
<td>Group 4</td>
<td>2 11</td>
<td>2 14</td>
<td>0 00</td>
<td></td>
</tr>
<tr>
<td>Group 5</td>
<td>1 6</td>
<td>1 7</td>
<td>0 00</td>
<td></td>
</tr>
</tbody>
</table>
The full professors and assistant professors were more likely than associate professors to name instruction as their primary activity area. Full professors and assistant professors were also more likely than associates to name research as their primary activity area. The most common activity profile for associate professors was a relatively balanced split between research and instructional activities. On the other hand, there were no assistant professors who reported a balance between research and instructional time. The assistant professors in this sample reported spending either a dominant amount of time in instruction or a dominant amount of time in research.

An important factor regarding the situations of these 41 faculty is the enormous range of activity profiles both within and across departments and across established activity profile groups. In other words, faculty assignments are very individualized. There is no single faculty profile from which the faculty role can be understood or viewed.

There is also great variety represented in what faculty do within an activity area. Some faculty teach small graduate courses, some teach very large enrollment undergraduate courses, and some a combination of both. Some faculty teach only one quarter a year, and others teach four quarters. Some faculty teach practicum-oriented courses which require different management and organizational time than classroom-based courses which focus on theory. Some faculty teach all laboratory courses, some teach only classes with lectures.

Regardless of what, how much, or how often particular faculty reported teaching particular courses, they described themselves as the parties responsible for and in control of choosing subject matter, teaching
methodologies, testing procedures, etc. They described the faculty as the instructional decision-makers within this setting. Further, individual courses are often designed, conducted and evaluated by individual faculty without any functional input from others. Faculty gave the following kinds of responses to questions such as: Has anyone else ever been involved in evaluating your courses? or What did you do when, and who was involved in, developing your courses?

I can do absolutely anything I want to do. No one really monitors what I'm doing, what I'm teaching or how I'm teaching it. I'm really on my own. Once the course is approved I can pretty much do what I want. (f06)

8-- is a course that was on the books for a long time but hadn't been taught. I think the last time it was taught before I came here was nine years ago. So there were no notes, no exams to look at. I had to do it entirely from scratch....It was a matter of pulling things from courses that I had that related at least somewhat to that course. So I used some of the notes that I had as a graduate student. Other than that it was mostly pulling together things from journal articles that I read that related to the subject. (f08)

Once you have developed a course, you have students, it pretty well goes on as your course and it continues on until you change it. (f11)

I was given a course outline and told there you go. (f11)

I try to keep abreast of the research material that comes out. For the last three years, I've had all of my notes on my microcomputer. So every year I do a rewrite by just updating the sections. My notes are essentially like a book, but it's printed from my micro....I just run copies and give them to my students.

Q: Has that ever been reviewed by anyone other than yourself? A: No. (f18)
The course is pretty much my own. I put it together and if people ask me for the outline I don't have any objections to doing that. I give it to them. But other than that I've never had a discussion as to what was in there. (f22)

When the course was developed it had absolutely no input. It was up to me. It was what did I think our students in this department needed. (f24)

First I thought there was a course that needed to be taught in our department but wasn't offered when I came here. I thought I should teach something that complemented what was already taught in the unit. I went to the faculty who were teaching the lower level courses. Each of them said: Do what you want to do and don't worry about us. I had absolutely no direction. For the first two or three years it was just a terrible process because I didn't know...what the student's backgrounds were or what they were getting in other courses. The other faculty essentially said you do your own thing and it will work out all right. Just teach what you think is right. It was a matter of don't bother me with this. You work it out, and I don't care. Whatever you want to teach is all right with me. The chair thought it was a good idea to talk to the other faculty, but I never shared their reaction with the chair. Periodically our intent is to talk about courses and how they fit together, but the actual realism has never materialized. (f25)

The participants in the interviews reported that as a faculty they sometimes discuss departmental instructional issues and concerns but that these discussions are usually episodic and informal and deal with general issues such as departmental graduation requirements or the possible need for a new course.

We do as a faculty occasionally look at courses and course content but nothing on a regular basis.

Q: When you do that, how does the process operate?
A: To my knowledge it has happened twice in the 11 years that I have been here. Two years ago we were looking at what courses our department majors should have. Another time we
were looking at what was offered in a particular major.... There is no systematic method where we evaluate a course every two years or anything. When we were discussing what our majors needed, we did it because advisors were talking to each other over lunch and we realized that some of our students were not taking a broad enough program of study. (f11)

These discussions among faculty were described as usually focusing on whether or not courses were repeating the same content and whether or not new courses were needed or other courses could be dropped.

The faculty responsible for a course series met because we became aware that in many of the courses we were introducing the same things. Some things were being introduced at least three times in the sequence. Redundancy was the single biggest issue at hand. (f03)

For the most part, the faculty who participated in the interviews reported little formal or informal contact with their colleagues or chairs related to their instructional activities. Sixty-eight percent of the sample said their chairs had never been in their classrooms. Of the remaining 32% (n=13), only two reported that they had discussed their instructional practices in any depth with their chairs. Both of these faculty were in an education-related discipline. Typical of the comments made by the 68% was the statement of an assistant professor:

We only have faculty meetings about once a quarter and when we do we don’t really discuss the curriculum. We don’t ever have a seminar or retreat on teaching. There’s very little interaction among faculty. My interactions during the day are with students and with people on the phone. Literally people can be in this building and in this hallway for three or four days and I’ll never have an occasion to interact with them. I interact with my graduate students, people out in the state, other students, etc. If faculty do talk, it’s pretty much about the basketball game. You just don’t interact that much. (f41)
Two more senior faculty commented on this lack of involvement in the following ways:

I have never been asked about my teaching. I'm talking about faculty meetings, college meetings, etc. I've never been asked about teaching. Isn't that strange in a university? If the alumni are coming in they ask what are you researching that is important? At faculty meetings someone might say: Stand up and tell us about your research program. But we don't have time for the teaching. At the college level they never ask about teaching. (f32)

I'm presently chair of our curriculum committee, and two meetings ago I brought out the fact that: Why don't we just try to get away from here and talk about our courses? What we're teaching and how we're doing it. In a serious way let's see if we need to do things differently. They all said it would never work. They said no one would show up because of lack of interest. "The system's working. Don't fight the system. Why fix something that ain't broke?" (f25)

The interviews with these faculty created an image of the university as an apiary, with the colleges as separate hives within which complex sets of activities occur and are needed in order to accomplish a diverse set of goals. Although each faculty member may be enacting the faculty role in a slightly different way than his or her colleague in the next office, each individual contributes an important and necessary function to their department, college and university as a whole.

These faculty projected the distinct impression that they are sincerely interested in providing quality instruction and that they view instruction as central to the university's missions. There were some who appeared more invested in their role as teacher than others, but there were no faculty who appeared unconcerned about the goal of providing quality instruction to
university students. Representative of these faculty members' attitudes toward the teaching role are the following faculty comments:

In my life, teaching has been important. Although I enjoy research, I never could spend my entire time in research. The classroom, the students— that's what is fun. (f02)

Personally, I think that teaching is what makes me feel good as a person. It is what gives me a sense of social accomplishment or something that is socially constructive. I'm sure you know you get feedback from students that you can't put into a form or responses to an SET form. I guess to some degree the research is a means to an end— to be able to teach. (f04)

I guess if I would have to say what my real driving force is it would probably be research, because I like thinking up new ideas and working with them. But in the same sense I do like to give a decent class, and I think one of our most important jobs is turning out graduate students. So I work in that area. (f22)
Section B: Formal Instructional Evaluation

Formal Instructional Evaluation: means through which information about instruction-related activities is purposefully collected, interpreted and used to make judgments and subsequent decisions about the selection, retention, maintenance or changes in instructional programs, personnel, curricula, courses, methods or specific learning activities.

Each of the 41 faculty interviewed in this study were asked to describe any means of instructional evaluation which they or others had used to evaluate their instructional practices. Whenever they mentioned a particular type of instructional evaluation procedure they were asked to describe the actual process in detail—including when it occurred, how often, the nature of the information, and perceived value of the information for making informed adaptations and improvements.

Fourteen different formal instructional evaluation methods were mentioned across these 41 interviews. Table 8 lists, in rank order, the types and reported use frequency of each type of formal instructional evaluation process as reported by the faculty in the sample. [NOTE: Written student evaluations refers to any type of formal student evaluation of instruction that includes written documentation, i.e., ratings instruments, open-ended questions, etc.]
<table>
<thead>
<tr>
<th>Type of Evaluation Process</th>
<th>Faculty Users</th>
<th></th>
<th>Users' Academic Units</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>% of Sample</td>
<td>n</td>
<td>% of Sample</td>
</tr>
<tr>
<td>Promotion and tenure reviews</td>
<td>41</td>
<td>100</td>
<td>11</td>
<td>100</td>
</tr>
<tr>
<td>Written student evaluations of instruction</td>
<td>33</td>
<td>80</td>
<td>10</td>
<td>91</td>
</tr>
<tr>
<td>Annual performance reviews</td>
<td>13</td>
<td>32</td>
<td>8</td>
<td>73</td>
</tr>
<tr>
<td>Analysis of student performance on tests</td>
<td>10</td>
<td>24</td>
<td>8</td>
<td>73</td>
</tr>
<tr>
<td>Departmental interviews/surveys of graduating seniors</td>
<td>10</td>
<td>24</td>
<td>6</td>
<td>55</td>
</tr>
<tr>
<td>Discussions with individual faculty for purpose of critique</td>
<td>6</td>
<td>15</td>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td>Group discussions with teaching faculty for purpose of critique</td>
<td>6</td>
<td>15</td>
<td>2</td>
<td>27</td>
</tr>
<tr>
<td>In-class observations by peers</td>
<td>5</td>
<td>12</td>
<td>4</td>
<td>36</td>
</tr>
<tr>
<td>Curriculum review committees</td>
<td>4</td>
<td>10</td>
<td>4</td>
<td>36</td>
</tr>
<tr>
<td>Individually requested review of teaching materials</td>
<td>4</td>
<td>10</td>
<td>2</td>
<td>27</td>
</tr>
<tr>
<td>Discussions with individual students or groups of students for purpose of course critique</td>
<td>2</td>
<td>5</td>
<td>3</td>
<td>27</td>
</tr>
<tr>
<td>Consultation with a teaching improvement consultant</td>
<td>2</td>
<td>5</td>
<td>2</td>
<td>18</td>
</tr>
<tr>
<td>Faculty development committee</td>
<td>2</td>
<td>5</td>
<td>2</td>
<td>18</td>
</tr>
<tr>
<td>Colleague rating form</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>9</td>
</tr>
</tbody>
</table>
Only three processes were mentioned by 25% or more of the interviewees: promotion and tenure reviews, written student evaluations of teaching, and annual performance reviews. The majority of formal methods were each reported by 15% or less of the sample. Table 9 lists the faculty members and indicates which methods were described by each.
<table>
<thead>
<tr>
<th>Promotion and tenure reviews</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Written student evaluations</td>
<td></td>
</tr>
<tr>
<td>Annual performance reviews</td>
<td></td>
</tr>
<tr>
<td>Written student evaluations</td>
<td></td>
</tr>
<tr>
<td>Promotion and tenure reviews</td>
<td></td>
</tr>
<tr>
<td>Faculty participation</td>
<td></td>
</tr>
<tr>
<td>Faculty development committee</td>
<td></td>
</tr>
<tr>
<td>Colleague rating form</td>
<td></td>
</tr>
<tr>
<td>Consultation with a teaching improvement consultant</td>
<td></td>
</tr>
<tr>
<td>Discussions with individuals or groups of students for purposes of critique</td>
<td></td>
</tr>
<tr>
<td>Individual requested review of teaching materials</td>
<td></td>
</tr>
<tr>
<td>Group discussions with teaching faculty for purpose of critique</td>
<td></td>
</tr>
<tr>
<td>Curriculum review committees</td>
<td></td>
</tr>
<tr>
<td>Departmental/intermediate surveys of graduating seniors</td>
<td></td>
</tr>
<tr>
<td>Analysis of student performance on tests</td>
<td></td>
</tr>
<tr>
<td>Annual performance</td>
<td></td>
</tr>
<tr>
<td>Written student evaluations</td>
<td></td>
</tr>
</tbody>
</table>
Promotion and tenure reviews, annual performance reviews, and written student evaluations of teaching were the most frequently mentioned processes through which instruction was reported to be formally evaluated. Promotion and tenure reviews and annual performance reviews are not separate evaluation methods per se, but rather processes through which information about instructional activities is often collected, usually through written student evaluations, and that information is then used to make personnel decisions. The faculty, however, seemed to pair "instructional evaluation" with promotion and tenure reviews and annual performance reviews in their discussions and, therefore, these reviews are treated in the findings as separate, formal instructional evaluation processes.

No individual faculty member reported having used all 14 different kinds of formal instructional evaluation. Instead, there were 32 different combinations of instructional evaluation methods reported by these 41 faculty. It is also important to note that, except for written student evaluations of teaching, when the faculty referred to having used a formal instructional evaluation method, they were usually describing something that had occurred only episodically during their faculty career rather than systematically. In most cases, they were describing an event that had occurred only once or twice in the history of their faculty career.

Table 10 lists each formal instructional evaluation process according to reported regularity of occurrence.
TABLE 10

Regularity of Occurrence

<table>
<thead>
<tr>
<th>Regularity</th>
<th>Evaluation Process</th>
<th>n</th>
<th>% of Sample Reporting Usage</th>
<th>% of Non-users</th>
</tr>
</thead>
<tbody>
<tr>
<td>Continuous (At least once/quarter)</td>
<td>Written student evaluations</td>
<td>33</td>
<td>80</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Self-analysis of student test</td>
<td>10</td>
<td>24</td>
<td>76</td>
</tr>
<tr>
<td></td>
<td>performance</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Discussions with individual</td>
<td>8</td>
<td>15</td>
<td>85</td>
</tr>
<tr>
<td></td>
<td>faculty for purpose of critique</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Episodic (Once or twice in history of teaching career)</td>
<td>In-class observations by peers</td>
<td>5</td>
<td>12</td>
<td>88</td>
</tr>
<tr>
<td></td>
<td>Curriculum review committees</td>
<td>4</td>
<td>10</td>
<td>90</td>
</tr>
<tr>
<td></td>
<td>Group discussions with faculty</td>
<td>4</td>
<td>10</td>
<td>90</td>
</tr>
<tr>
<td></td>
<td>for purpose of critique</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Materials review</td>
<td>4</td>
<td>10</td>
<td>90</td>
</tr>
<tr>
<td></td>
<td>Discussions with students for</td>
<td>2</td>
<td>5</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td>purpose of critique</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Consultation with teaching</td>
<td>2</td>
<td>5</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td>improvement consultant</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Faculty development committee</td>
<td>2</td>
<td>5</td>
<td>95</td>
</tr>
<tr>
<td>Once per year</td>
<td>Annual reviews</td>
<td>13</td>
<td>32</td>
<td>78</td>
</tr>
<tr>
<td></td>
<td>Exit interviews</td>
<td>10</td>
<td>24</td>
<td>76</td>
</tr>
<tr>
<td></td>
<td>Colleague rating form</td>
<td>1</td>
<td>2</td>
<td>98</td>
</tr>
<tr>
<td>Once in 5-7 years</td>
<td>Promotion and tenure reviews</td>
<td>41</td>
<td>100</td>
<td>--</td>
</tr>
</tbody>
</table>
According to these interviews:

- 15% of the sample reported conducting no ongoing formal instructional evaluations.

- 51% reported written student evaluations of teaching as their only method of conducting ongoing formal instructional evaluation.

- 5% reported self-analysis of student performance on tests as their only ongoing formal instructional evaluation process.

- 10% reported written student evaluations and self-analysis of student test performance as their only ongoing means of conducting formal instructional evaluations.

- 12% of the sample reported using written student evaluations and discussions with individual faculty for the purpose of critique as ongoing formal instructional evaluations.

- 7% reported using written student evaluations, self-analysis of student test performance, and discussions with individual faculty for the purpose of critique as ongoing formal means of evaluating their instructional practices.

There is not a one-to-one correspondence between teaching load index or faculty rank and pattern of instructional evaluation practices. In general though, those faculty with the highest teaching loads tended to report having used a variety of instructional evaluation methods throughout their careers, while those with lower teaching load indices tended to report less involvement in instructional evaluation as a formal procedure. In addition, faculty rank appears to be related to instructional evaluation practices in that 61% (n=11) of the full professors described involvement in only three or less different kinds of instructional evaluation activities over the span of their career. In contrast, 57% of associates, and 67% of the assistant
professors described involvement in four or more different instructional evaluation processes over the span of their careers.

Nine of the 14 methods mentioned by these faculty have been described in the literature as particularly useful for improvement purposes. These are: written student evaluations, analysis of student test performance, discussions with individual faculty for purpose of critique, in-class observations by peers, group discussions with faculty for purpose of critique, materials review, discussions with students for purpose of critique, consultation with a teaching improvement consultant, and faculty development committees. Except for written student evaluations of teaching and self-analysis of student test performance, none of these nine methods were reported by more than 15% of the sample, and the great majority were reported by 10% of the sample or less. This leads to the conclusion that, in general, formal instructional evaluation methods are not, to any large degree, being used by these faculty as a means of critiquing their instructional practices for purposes of improvement and development.
Promotion and Tenure Reviews

All of the full and associate professors cited promotion and tenure reviews as a process through which their instructional performance had been evaluated. The assistant professors cited future review as a process for which they were collecting information about their teaching and through which they expected their instructional performance to be evaluated.

There was a marked difference between the ways full professors who had been on campus ten to 30 years and full and associate professors who had been on campus ten years or less described their personal experiences with promotion and tenure reviews. The oldest professors said they had played no role in documenting their teaching for promotion and tenure, and they had no idea exactly what information was considered. They usually learned they received promotion and tenure when they discovered a salary increase and title change on their new contract. Three of the more senior faculty described their experience in the following ways:

When I went through the tenure process, I don’t recall submitting any information about my teaching....I didn’t even know my name was going forward to be considered--I didn’t know it until I learned from the chairman that I had been tenured. We were asked each year to bring our CVs up to date, and I did that. And he talked to graduating students about courses. I assume that he obtained most of his information that way. I don’t know to what extent I was evaluated on my teaching. I expect it was mostly on the research I had done. I don’t recall ever getting any written communication about it until I actually got my contract to sign, and it came with Associate Professor on it. (f18)

Promotion and tenure is a very different thing than when I first came here. I recall being in the lab and the department chair
came in and I nonchalantly said something like: By the way, do we get tenure or what happens? He said: You got tenure two years ago. Today you'd know it. I'd been tenured for two years and didn't know it. He had sent it in. He didn't ask any committee or inform me. There was no teaching instrument involved. I doubt if I even submitted it for full professor. (f02)

I remember back in '76 or '77 there was a big-to-do about teaching evaluation. You had to have teaching forms. It wasn't uniform or widespread. It was sporadic. People in the past didn't have to be concerned about this. They were able to down play it. From the mid-seventies on it was necessary to carefully document. (f12)

I came on board back in the time when tenure was not as difficult as it is today. So I made no special efforts to build up a vitae that would be looked upon favorably from a promotion or tenure or salary standpoint...I could not tell you when I was promoted to full professor. It was nice, but it wasn't a milestone. (f28)

On the other hand, the full and associate professors who had received promotions and/or tenure in the last ten years said that they had been required to provide reviewers more and more detailed dossiers, including evidence about their teaching. In most cases the evidence provided was restricted to summaries of student evaluation results and a description of their courses according to enrollment, credit hours, etc. They also included any course development activities they might have been involved with.

I included some kind of a summary of my SET [Student Evaluations of Teaching] evaluations. I included new initiatives, changes in courses. I can't think of anything else. (f10)

What I did was take my student comments. I randomly selected four or five comments and wrote them up. (f11)
I list the courses I’ve taught. The numbers. Also the SETs. Also any new courses I’ve developed. (f16)

The mean scores on the six core items of the SET. And sort of pull together information that was germane to the course. (f17)

I put together the things that I had improved in tremendously and what my percentile rankings were on those.

Q: Did you only select those things that you had improved on?
A: Yes...Look, if they want to find the negative, they can find it. I think I had the whole thing appended. But they say: If somebody writes a letter and says one little small thing about you that’s even shady, they don’t put it in. Promotion and tenure is how you’ve improved. (f42)

When P&T was first discussed with me, I picked out a whole pile of evaluations. My chair told me to pick out a couple, summarize or do something because they didn’t want to see all of those...I was a little offended by that. (f06)

I put together some materials—but these packets have grown and grown since I had to do it. It wasn’t really very clear what should be put into the packet. The rules keep changing and you need more and more documentation. The rules seem to be that you should put in everything. (f06)

I keep evaluations of public service activities. Thank you notes from people I’ve conducted clinics for. I have all my datebooks, numbers who attended, etc. I have the student evaluations of my course for every course I’ve taught. (f21)

We put together a packet based on our annual reviews and everything else we’ve constructed. I have a file with a section called yearly accomplishments and it’s really good because every time I do anything good I stick it in there. For teaching accomplishments I have my lab manual in there—-I think that’s good—and when my book comes out I’ll put that in there. The SETs are filed for documentation. (f25)
The second time I took more of a leadership role in putting together the documentation. The first year I just turned in stuff—lists of publications, lists of meetings, letters from people or names of people that could write letters in support of my programs—that kind of stuff. The second year I not only turned in those lists of publications, but I wrote the narrative describing the nature of the lists of publications. I took a much stronger leadership role because points were being missed.

(f31)

Only four faculty described any process through which colleagues observed their in-class teaching performance as part of the promotion and tenure review; in each of these instances the faculty received no specific feedback regarding these persons’ observations. Written reports were apparently given to chairs and others involved in the review. Faculty remarked that they assumed they performed satisfactorily because they never heard otherwise.

I was observed...once by a faculty member who was on a committee assigned to observe me when I was first up for promotion and tenure. He and the chair were the only two from a committee of four who actually visited by class. I guess they used that for P&T. (f06)

There was a committee set up to evaluate my teaching. One guy came to ten classes, the department chair came once...and then we never met. I guess they decided I was doing O.K. I don't know otherwise. (f24)

Professor ---- was in my class and I asked him to stay...and write up a report on how things went which he did and then he shared that with my chair for my file. (f27)

Our past chair recommended to new faculty or to anyone teaching to avail themselves of peer review. You don't see a lot of it happening. I asked a professor who was interested in the subject matter of my course and therefore sat in on it to write up an evaluation. He did that and gave it to the chair. (f40)
one department was reported to have peer review committees which are available to be used by untenured assistant professors and whose function was described as "to help young faculty move toward promotion and tenure (f26)."

It is a mentoring relationship....We review, not only their teaching function with them, but their service and their research activities too. (f26)

Other than these voluntary peer review committees in one department, no faculty reported review by colleagues, chairs, or others of the quality of subject matter, exam practices, student outcome measures, or instructional materials as part of promotion and tenure reviews. They did report that letters by colleagues, chairs, and persons outside of the university were often included in their promotion and tenure dossiers. Since these were reportedly usually added after the professor submitted his materials to the chair, they did not serve as a direct feedback source to the faculty member. Because the supporting letters of these people were reportedly so seldom based on first-hand review of teaching performance or courses, any references to the faculty member's instruction-related performance were assumed by the interviewees to be derived from general impressions or hearsay from students.

The feedback received in the last ten years by faculty as part of promotion and tenure review was reportedly a general letter of acceptance or rejection. They reportedly interpreted their acceptance as meaning they were generally performing satisfactorily, but they were given no specific feedback about separate aspects of their performance.
One way that faculty use promotion and tenure reviews as performance feedback is by noticing who is accepted and who is rejected along the way. They interpret these personnel decisions and use the information for directing and documenting their own performance. It appears to them that excellent researchers are more likely to gain promotion and tenure than excellent teachers, even though "no one will openly admit it (f25)."

The faculty's general impression is that evidence of excellent teaching performance without satisfactory research accomplishments is not sufficient for promotion or tenure, but excellent research accomplishments without satisfactory teaching performance are sufficient. None of the participants was able to describe specific university or department standards for excellent or satisfactory performance in either area. They typically commented that they thought satisfactory teaching performance was measured by "not too many students complaining to the chair."

What's rewarded is research, publications, grants. I've been involved in both. Teaching has been rewarded for me for its own sake. I don't think my teaching ability has really been considered in terms of promotion, salary, whatever. We see this with colleagues at P&T. What they want to know is how many papers have you published? How many grants? (f01)

I think in the 60s, during the student uprisings, I think we attained a level where all administrators deemed it proper to say we'd evaluate teaching as much as research. Well, I never believed them, but I found it enjoyable to hear them say that. I guess we just have to accept the fact that in an institution of this size, a teacher can never have the value that a good researcher does. (f02)

I don't know how it's weighted. I guess I don't think there are very many rewards for being a good teacher. I don't think quality teaching is recognized. I don't think bad teaching is
recognized. You can be a pretty rotten teacher at this university and get away with it. Our dean said that there is room for one "super" teacher in our unit. That gives me an indication of how teaching is viewed. (f05)

It's just something you know. Research and grant money are valued higher than teaching efforts. The message comes from the university and college level. The department sort of transmits those kinds of things. (f06)

As far as P&T is concerned, I would say it probably isn't very clear to me on what criteria my teaching will be judged. Occasionally we have P&T meetings within the college. Where they invite the dean and other administrators and they present their feelings on P&T and the faculty can ask questions about things. The administrators nearly always say: Yes, we do put a lot of emphasis on teaching. It's very important. It's going to count heavily in your P&T. I think the general consensus is it comes down primarily to how many publications you've got. Whether that is really what happens, I don't know. But I guess my feeling would be that if you're a good researcher with a lot of publications and have brought in a lot of grants, but you're not a very good teacher, they'll probably still figure out a way to keep you around. Whereas, if you're a super teacher but have done very little research, they may not keep you around. (f08)

When I had a postdoc at a different university, the general consensus in the department was that one of the professors who was the best teacher in the department didn't get tenure. You hear stories like that and you kind of tend to form the opinion that they don't really look at teaching all that much when it comes right down to it. Perhaps if you're really a terrible teacher...if your evaluations are extremely poor they would consider that. But as long as you're at least kind of an average teacher and doing a pretty good job on research, at least my perception is, you'll get by. (f08)

I would suspect that a lot of effort went into convincing me and others that were up for tenure that teaching was important. But in the end I suspect it was the research publications that did it. (f12)
I feel that my promotion and tenure decisions were made directly on my research productivity rather than as an educator. (f15)

I think unless you are a teacher that is internationally recognized that it is very difficult for teaching to help you, but it is possible that if you are a really poor teacher—to the extent that the students have gone to your department chairman or your dean and said this person is absolutely miserable—that teaching can hurt you. But if they went to my department chairman and said: Dr. ---- does a really nice job teaching. He works us really hard, but I've learned a lot in his classes. That's what you are supposed to do. So I don't think that would help you any except in a very rare instance. (f20)

If I didn't have publications I'd be sunk—I don't care how good my teaching is—even though I formally have a 75% teaching assignment. (f25)

Everyone who teaches must do research of a certain caliber. When was the last time you heard someone say everyone who does research must teach at a certain level of excellence? (f27)

I think bad teaching may hurt you. I don't think good teaching can help you. It's not really a very important function here. Even though lip service is paid to it. I think from what I've seen that good young people with good teaching awards don't make the centerfold of the football program. They are bounced out because of research. I've never seen the converse unless it was a morals charge or something. Somebody who's doing a very good job of research and a very bad job of teaching, I don't think would have any trouble. Somebody doing a good job of teaching and a bad job of research I think is in serious trouble. (f32)

I contend that I could stay here for the next two years and not teach course one and not do public lecture one. As long as I cranked out the articles I could get promoted and get tenure. The only reasons that teaching is listed as part of the P&T
process is because the university does teaching, service, and research. It is bandied about that we do teaching so obviously you have to include it. Nobody ever looks at it. You see people go in with what I would consider deplorable evaluations, but they have a list of publications the length of your arm and they pass with no problem. (f34)

From everything I've heard—from listening to the provost and talk—all you have to do is demonstrate that you are a sufficient teacher. That you are taking it seriously. It is a kind of mini/max thing. You don't want to spend all of your time trying to become the best teacher. What you are doing is you are always looking at the margin. What are some of the things that I can get the greatest return on for the least amount of effort to put in? (f40)

We talk about how we fared and about how others fared. You know how the gossip goes. This guy made it because he's writing, but the guy stinks in teaching. And you get information from students about how other professors teach. And sometimes they can't understand how that guy made tenure when you're a better teacher—but he's a better writer. (f42)

Standards of satisfactory research performance were described by some as one article in a refereed journal each year. They were described by others as five such publications a year. Still others described publications in refereed journals and research grants as necessary for promotion to full professor. Most of the faculty reported that the expectations for satisfactory performance are undefined and unclear. The associate and assistant professors almost invariably expressed a desire for more explicit definition of performance standards and expectations so that they could understand more clearly how to direct their actions. The faculty seemed to agree that more and more documentation is being required for promotion and tenure and that the individual faculty member is primarily responsible for assembling the dossier. Albeit, once the dossier is submitted the faculty
were reportedly unsure of how or if information about teaching was specifically treated, and they reported receiving no feedback other than a final decision.
Written Student Evaluations

Thirty-three of the 41 faculty participants claimed that they currently conduct written student evaluations of their courses. The claims of these 33 were confirmed with physical samples they provided. Three additional faculty participants claimed to collect written evaluations from students, but their claim could not be substantiated with physical evidence. Based on this lack of evidence, confirmed in their interviews, it was judged that these three faculty did not actually (currently) conduct written student evaluations, and their responses were analyzed assuming they were non-users. The remaining five participants reported that they did not conduct written student evaluations of their courses.

The physical samples provided by the 33 current users were analyzed as to type of format used, and the faculty were grouped accordingly. The comments for each group were reviewed as to reported reasons for administering the evaluation, claimed value of the process, methods of evaluating the findings, reported changes made as a result, and factors which seemed related to specific practices.

Formats in Use

As a group, these faculty used a variety of student evaluation instruments and items. The varied formats can be classified into essentially three main categories: 1) formats with primarily rating items (some of which include a limited number of general-comments items); 2) formats with descriptive items only; and 3) formats which include a substantial number of both rating and descriptive items. Table 11 summarizes the usage of these various student evaluations.
### TABLE 11
Types of Instruments and Formats Used for Written Student Evaluations

<table>
<thead>
<tr>
<th>Instrument/Format</th>
<th>n</th>
<th>% of Users</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primarily rating items</td>
<td></td>
<td></td>
</tr>
<tr>
<td>University SET</td>
<td>9</td>
<td>27</td>
</tr>
<tr>
<td>Cornell</td>
<td>7</td>
<td>21</td>
</tr>
<tr>
<td>Non-standardized</td>
<td>5</td>
<td>15</td>
</tr>
<tr>
<td>Subtotal</td>
<td>21</td>
<td>64</td>
</tr>
<tr>
<td>Descriptive items only</td>
<td>7</td>
<td>21</td>
</tr>
<tr>
<td>Combination of formats</td>
<td>5</td>
<td>15</td>
</tr>
<tr>
<td>Total</td>
<td>33</td>
<td>100</td>
</tr>
</tbody>
</table>

**Rating items** are defined as those which request that a numerical value be assigned to some designated dimension of instructor behavior or the course. Rating items were used as the primary data collection tool by 21 of the 33 faculty. Two types of rating items instruments were used, standardized and non-standardized. In addition, two standardized instruments were used: 1) a university form (SET), and 2) the Cornell Diagnostic and Observation Instrument.
The University SET is designed so the faculty member can select up to 20 rating items (a 5-point rating scale) from a menu of 200+ items. The instrument also includes five required "core items" selected by the university administration and intended to provide a set of common questions of high generalizability that are applicable to all, or nearly all, courses. In addition, the faculty member can ask the students to provide written comments, or responses to open-ended questions, on the back of the response form. The rating items are machine-scored, and a statistical summary of individual, college and university-wide results are returned to the faculty member along with the response sheets. Analysis and interpretations of the students' written comments are left up to the instructor.

The Cornell Diagnostic and Observation Instrument is a student evaluation of teaching tool that is offered through the college. It was developed at Cornell University and norm-referenced using "exemplar" college teachers as a norming group. It includes 28 rating items (5-point scale) which focus on specific in-class teacher behaviors. It includes an additional seven items in which the student can rate how effective he or she thinks the instructor was at accomplishing seven general teaching objectives. The instrument is machine-scored and a computer printout of individual item-by-item and norm-referenced comparative statistics is provided to the instructor. The form does not allow for open-ended items or for selecting items.

Descriptive items are defined as those which ask the student to describe his or her experience or opinion of certain aspects of the course, or simply general impressions. They require no numerical value to be assigned as a
response. Descriptive formats were used by seven participants and ranged from extremely unstructured to extensive lists of course-specific, open-ended items. The least structured of these included one in which a faculty member asked the students to take out a blank piece of paper and comment on the course:

[I tell them to] think about the course--the lecture, the labs, the instructor, the TA. I suggest categories for them to comment on, which usually differ from course to course because my courses differ in organization and structure--some have labs, some don't etc. Or they can make comments outside of these categories. I give them a pretty free rein. (f28)

Another faculty member lists several open-ended questions on the board and asks the students to write their responses on a piece of paper. These questions include asking them to address such things as:

The depth versus the breadth of the subject matter, the usefulness and relevancy of the text, the examinations, their opinions of me as a instructor--Do I appear prepared for class? Am I interested in the students? Do I know the subject matter? Do I keep them interested? Do I have any annoying habits? Or anything else they'd like to comment on regarding the course. (f11)

Five faculty members from the sample used a combination of written student evaluation formats. Those formats included either more than one type of form for the same class or a single form with a substantial number of both rating and descriptive items. Three faculty used both a standardized rating items form and a self-designed form with descriptive items for the same class. The other two used a self-designed form that included a substantial number of both non-standardized rating items and descriptive items.
The relationship between instructional improvement and development efforts and written student evaluations varied by the type of instrument used. Rating items were used primarily for promotion and tenure purposes and were viewed as minimally useful for making improvement and development choices, while descriptive items were viewed as particularly useful for improvement and development aims.

**Reasons for Using Written Student Evaluations**

Clearly, the large majority of the users of written student evaluations administer rating items as a means of collecting the information (64%). However, only two of the users of rating items reported improvement as their primary reason for conducting such evaluations. Four of the 21 users of rating items were full professors who were either required by their department to administer a group of 15 department-selected SET items, or who had been instrumental in getting the Cornell to be used in their college and felt an obligation to continue using it. The remaining 15 users (71%) cited documentation for promotion and tenure as their primary reason.

On the other hand, no one who used only descriptive items reported promotion and tenure as their primary reason for conducting student evaluations of their courses. Instead they all defined their guiding rationale in terms that referred to:

1) A desire to acquire information that might help them continue to develop and improve their courses;

2) The perception that written comments and opinions, rather than numerical values assigned to predefined dimensions, are more adequate for collecting information useful for identifying improvement needs.
The users of a combination of formats tended to use rating items to provide numbers or comparative data for administrators, but administered descriptive items for the purpose of self- and course improvement. All five had high teaching loads, representing five of the ten highest teaching loads across the sample. All five stated or implied a particularly strong commitment to continued self-improvement and development.

Table 12 displays the faculty's remarks pertaining to their reasons for using written student evaluations of teaching.
Our chairman asks us to include a summary of the SET results with our annual report packets.

We decided several years ago as a department we would use this uniform set of questions and the results would be shared with the chair.

Because I think I ought to have some way of communicating how I'm doing to my chairman.

The department requires it. I need it for documentation of teaching for promotion and tenure.

Because the provost says you have to do them.

The department requires it. To file for promotion and annual reviews.

They'll go in the file, basically for promotion and tenure.

To have documentation for promotion and tenure. You need a section on student evaluations in your promotion and tenure packet.

-I used to use the SET. I got cynical. I didn't think it told me anything. It wasn't bad or anything. I always came up above average. I just didn't get much feedback. Then from the SET I went to the Cornell, I kind of liked the Cornell and I got good feedback on that. But I never got feedback from the students in terms of them telling me. They would circle an answer. It seemed that if the students had to write a comment it was more meaningful than circling an answer.

-In the lower classes I have used the university SET in the past. I have been most unhappy with them. It is treated almost like an evaluation of the instructor and not a way to help the instructor improve. This is not an instructor evaluation. You will not find any questions about the instructor. This form is an evaluation of the course.

-Four years ago I objected to the

-I find them very useful for self-evaluation. I also use them for promotion and tenure.

-Primarily the research that I have been doing has been revolving around teaching. I've done some extensive work on course evaluation over the last decade. The first quarter I was teaching I had an honors student looking for an honors project so he worked with me on an evaluation form and what he wanted to see was how students evaluated that course at the end of the quarter and how they looked at it one year later. Out of that we developed a questionnaire, an evaluation form that we thought was a pretty good one. I've collected data with the same instrument for every course I've taught for the last ten years and I've used the data to make some different kinds of comparisons. I've written some articles and made some presentations based on the findings. They help me find out what needs to be changed.

<table>
<thead>
<tr>
<th>TABLE 12</th>
<th>Reported Reasons for Using Written Student Evaluations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Users of Rating Items</td>
<td>n = 21</td>
</tr>
<tr>
<td>Users of Descriptive Items</td>
<td>n = 7</td>
</tr>
<tr>
<td>Users of a Combination of Formats</td>
<td>n = 5</td>
</tr>
</tbody>
</table>

-I used to use the SET. I got cynical. I didn't think it told me anything. It wasn't bad or anything. I always came up above average. I just didn't get much feedback.

In the lower classes I have used the university SET in the past. I have been most unhappy with them. It is treated almost like an evaluation of the instructor and not a way to help the instructor improve. This is not an instructor evaluation. You will not find any questions about the instructor. This form is an evaluation of the course.

-Four years ago I objected to the
Users of Rating Items

-Promotion and tenure—this is the reason... I have to have some way to document by abilities in the classroom.
-I attach them to my appendix in my annual report.
-The chair gets all that stuff to read. It gives him some numbers to look at. Of course the comparisons that are generated there he can look and see how I'm doing relative to how some other instructors in the college are doing.
-I use a format that was developed by another professor who teaches another section of the course I teach. I talked to him about it and he said he thought it had been useful for him for PAT so I used it—he did get tenure.
-Primarily to document for PAT
-I now use a format with a ten point scale on items because it's easier for the promotion process—easier than open-ended items.

Users of Descriptive Items

-University one where you filled in all these little spaces where you agree or don't agree. I try to give them one where they write a short paragraph if they wish as well as kind of evaluate the course, text, and the instructor. I'm still watching this. That's my way of checking—is there something going on here that I didn't know about? Then I can make some changes.
-Primarily because I think it channels the students' thoughts into areas that you are interested in getting responses rather than giving them total freedom to respond in areas that they think are critical. I don't want to lead them very much in an evaluation. I'm a full professor, tenured, and I use the information primarily to find out what I should be doing differently.

Users of a Combination of Formats

-I'm still trying to make improvements on my ratings on the Cornell. I used the SET but I found that after using it twice I had what score I was going to get—I was in the 90th percentile on everything. I didn't feel it was serving any purpose other than providing something for the department chairman to put in my file. I've already got tenure and promotion, I really don't think that salary advances are based on what one does in an evaluation of teaching. I think the important thing is that we do them. So if you are really doing something bad it will be caught.
-I tend to use the evaluation to find out whether they think we're putting too much emphasis on some subject, or whether subject matter needs to be changed more, or we're missing something, what they like about the use of visual aids, etc.—particularly from the upper-level students. I think those students know what they should know and what they should be able to utilize when they leave.

---

Table 12
(Continued)

<table>
<thead>
<tr>
<th>Users of Rating Items</th>
<th>Users of Descriptive Items</th>
<th>Users of a Combination of Formats</th>
</tr>
</thead>
<tbody>
<tr>
<td>-University one where you filled in all these little spaces where you agree or don't agree. I try to give them one where they write a short paragraph if they wish as well as kind of evaluate the course, text, and the instructor. I'm still watching this. That's my way of checking—is there something going on here that I didn't know about? Then I can make some changes.</td>
<td>-I'm still trying to make improvements on my ratings on the Cornell. I used the SET but I found that after using it twice I had what score I was going to get—I was in the 90th percentile on everything. I didn't feel it was serving any purpose other than providing something for the department chairman to put in my file. I've already got tenure and promotion, I really don't think that salary advances are based on what one does in an evaluation of teaching. I think the important thing is that we do them. So if you are really doing something bad it will be caught.</td>
<td>-I tend to use the evaluation to find out whether they think we're putting too much emphasis on some subject, or whether subject matter needs to be changed more, or we're missing something, what they like about the use of visual aids, etc.—particularly from the upper-level students. I think those students know what they should know and what they should be able to utilize when they leave.</td>
</tr>
</tbody>
</table>
Users of Rating Items

- I heard a lot about the Cornell and
  got the Associate Dean of Instruction
to invite one of its designers in to do
a seminar on it.
- I was on the college teaching
committee and it began to push the
Cornell. I felt some obligation to
use it.
- I could have some solid feedback
give information to them, and if I do a good
job with them, the promotion should come. I
don't do things for promotion—I do things
to do a good job—that's my philosophy of life.
I keep these on file to use for promotion.
- I have an aversion to filling out questionnaires
and evaluation of what I was doing
and how to improve it.
- I try to do one thing or another
because I don't know what the re-
quirements are now for sure for
promotion and tenure.
- I recognized the need to have some
give information to them, and if I do a good
kind of evaluation for promotion
job with them, the promotion should come. I
and tenure.
don't do things for promotion—I do things
to do a good job—that's my philosophy of life.
- I have an aversion to filling out questionnaires
and evaluation of what I was doing
and how to improve it.
- I try to do one thing or another
because I don't know what the re-
quirements are now for sure for
promotion and tenure.
- I recognized the need to have some
give information to them, and if I do a good
kind of evaluation for promotion
job with them, the promotion should come. I
don't do things for promotion—I do things
to do a good job—that's my philosophy of life.
I keep these on file to use for promotion.
- I have an aversion to filling out questionnaires
and evaluation of what I was doing
and how to improve it.
- I try to do one thing or another
because I don't know what the re-
quirements are now for sure for
promotion and tenure.
- I recognized the need to have some
give information to them, and if I do a good
kind of evaluation for promotion
job with them, the promotion should come. I
don't do things for promotion—I do things
to do a good job—that's my philosophy of life.
I keep these on file to use for promotion.
- I have an aversion to filling out questionnaires
and evaluation of what I was doing
and how to improve it.
- I try to do one thing or another
because I don't know what the re-
quirements are now for sure for
promotion and tenure.
- I recognized the need to have some

Users of Descriptive Items

give information to them, and if I do a good
job with them, the promotion should come. I
don't do things for promotion—I do things
to do a good job—that's my philosophy of life.
I keep these on file to use for promotion.
- I have an aversion to filling out questionnaires
and evaluation of what I was doing
and how to improve it.
- I try to do one thing or another
because I don't know what the re-
quirements are now for sure for
promotion and tenure.
- I recognized the need to have some

give information to them, and if I do a good
job with them, the promotion should come. I
don't do things for promotion—I do things
to do a good job—that's my philosophy of life.
I keep these on file to use for promotion.
- I have an aversion to filling out questionnaires
and evaluation of what I was doing
and how to improve it.
- I try to do one thing or another
because I don't know what the re-
quirements are now for sure for
promotion and tenure.
- I recognized the need to have some

give information to them, and if I do a good
job with them, the promotion should come. I
don't do things for promotion—I do things
to do a good job—that's my philosophy of life.
I keep these on file to use for promotion.
- I have an aversion to filling out questionnaires
and evaluation of what I was doing
and how to improve it.
- I try to do one thing or another
because I don't know what the re-
quirements are now for sure for
promotion and tenure.
- I recognized the need to have some

Users of a Combination of Formats

- I use the SET and a self-designed form
for each course I teach. I want infor-
mation about teaching from any class I
teach. The SET gives me an overview and
a set of numbers I can give to my chair.
But it doesn't give me the depth that I
want to know about what I'm doing—not
only what I'm doing wrong, but what I'm
doing right too. I'm getting a ceiling
effect from the SET—I'm high on every-
thing—it was not really contributing to
my knowledge of what could be improved in
that class.

<table>
<thead>
<tr>
<th>TABLE 12</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Continued)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Users of Rating Items</th>
<th>Users of Descriptive Items</th>
<th>Users of a Combination of Formats</th>
</tr>
</thead>
<tbody>
<tr>
<td>- I heard a lot about the Cornell and got the Associate Dean of Instruction to invite one of its designers in to do a seminar on it.</td>
<td>give information to them, and if I do a good job with them, the promotion should come. I don't do things for promotion—I do things to do a good job—that's my philosophy of life. I keep these on file to use for promotion.</td>
<td>- I use the SET and a self-designed form for each course I teach. I want information about teaching from any class I teach. The SET gives me an overview and a set of numbers I can give to my chair. But it doesn't give me the depth that I want to know about what I'm doing—not only what I'm doing wrong, but what I'm doing right too. I'm getting a ceiling effect from the SET—I'm high on everything—it was not really contributing to my knowledge of what could be improved in that class.</td>
</tr>
<tr>
<td>- I was on the college teaching committee and it began to push the Cornell. I felt some obligation to use it.</td>
<td>- I have an aversion to filling out questionnaires and many times they require an opinion about everything that you may not have an opinion about. So when I get a questionnaire if I'm interested I'll write a short note about two or three of the topics that are interesting to me. I use the same approach in my class. I just ask a couple of questions and if I pick up one or two points that seem to be a consensus of the class then I can work on those and make an improvement if I need to.</td>
<td></td>
</tr>
<tr>
<td>- I could have some solid feedback and evaluation of what I was doing and how to improve it.</td>
<td>- I try to do one thing or another because I don't know what the requirements are now for sure for promotion and tenure.</td>
<td></td>
</tr>
<tr>
<td>- To provide information for promotion and tenure.</td>
<td>- I use the University SET. But I just come out with almost all fives. That doesn't help me. It helps me for promotion and tenure, because they're all just excellent. I come out the finest teacher they've ever had—but it doesn't tell me a thing about the course. I'm concerned with how to improve my instruction, and so I find my self-designed forms really valuable—to have the students talk about what they see as important in the course and what they don't.</td>
<td></td>
</tr>
<tr>
<td>- For promotion and tenure.</td>
<td>- I recognized the need to have some kind of evaluation for promotion and tenure.</td>
<td></td>
</tr>
</tbody>
</table>
PLEASE NOTE:

This page not included with original material. Filmed as received.

University Microfilms International
Perceived Value of Written Student Evaluations for Improvement Purposes

Even though the majority of the users of written student evaluations used primarily rating items, none of the users described rating items as useful for making instructional adaptations and improvements. The SET users were openly critical of the usefulness of the ratings data, and the users of the Cornell and the non-standardized rating items expressed their value in relationship to negative experiences they had had with the SET rather than positive regard for their rating formats. At the same time, the faculty participants agreed that descriptive items and written comments from students were useful for identifying the need for and making course changes and improvements. Table 13 displays the faculty's comments regarding the negative features of rating items and their comments regarding the positive value of descriptive items.
### TABLE 13

**Negative Comments Related to Rating Items**

<table>
<thead>
<tr>
<th>I've seen enough of these to know where I'm going to come out. Certainly the last five or six years I have come out about the same.</th>
</tr>
</thead>
<tbody>
<tr>
<td>After so many years of this, you know what is going to be there. That doesn't vary much.</td>
</tr>
<tr>
<td>Nothing has ever shown up that was a surprise. I think I can evaluate how well I've done without that.</td>
</tr>
<tr>
<td>I can tell you what the students are going to say before I hand it out to them.</td>
</tr>
<tr>
<td>I can usually predict what the students will comment.</td>
</tr>
<tr>
<td>I know the problems and sometimes I feel that I know the problems well enough to work on my own as compared to needing a formal evaluation.</td>
</tr>
<tr>
<td>I administer the SETs, but I don't think they are that valuable.</td>
</tr>
</tbody>
</table>

**Positive Comments Related to Descriptive Items**

<table>
<thead>
<tr>
<th>I have found the students' answers to my open-ended questions very useful and I've found that the students are very open on that kind of a thing.</th>
</tr>
</thead>
<tbody>
<tr>
<td>I try to convince the students that I take it seriously, but that I'm particularly interested--I feel I get the most from written comments. It's perhaps one out of five or six students that will address a point and say it in a way that you may have heard it before but said in a way that gets to you. Convinces you that you've got to work on that and do something about it.</td>
</tr>
<tr>
<td>I feel like if they (students) feel strongly enough to comment--positive or negative--everyone fills out the front, but not everyone takes the time to put down written comments. They carry more weight--whether it's formally or informally in my mind.</td>
</tr>
<tr>
<td>I get good ideas from the students. The kind I like the best is the open-ended. Those are the ones that I get the best feedback from.</td>
</tr>
<tr>
<td>Negative Comments Related to Rating Items</td>
</tr>
<tr>
<td>------------------------------------------</td>
</tr>
</tbody>
</table>
| I don't think the SET is worth the paper it's printed on. I don't think the Cornell is worth the paper it's printed on. I think the questions are trite. They call for pat responses from the students. I don't think they ask for an in-depth investigation of what really takes place in a classroom situation. I always tell my students: "This is an SET. I have the opportunity to pick from some 200+ questions those which are going to appear on this form. Therefore, it should be obvious to you who are intelligent individuals that I have self-selected questions which are going to be of most benefit to me."
| I usually turn first to the comments on the back. My lab technician and I go through them and consider their comments. |
| I used the SET once but wasn't satisfied with the information it provided. I tried the University SET when I first came here. It had questions like: The instructor usually comes to class on time, etc. I thought that was ridiculous. Then we had the Cornell form. It had just started. It took me a long time to get it back, and I wasn't sure that it was telling me anything. I want to a form basically that said: Tell me what you think about the course. There's a lot of good things about that. In fact, I like that very much. But it doesn't make the department chair very happy, and that type of thing. |
| What I think is more valuable [than ratings] is written comments. I stress with the students that I get more input and better feedback from the written comments than I do from coloring in the circles. |
| I look very carefully at the comments that students write and I try to put them into some perspective. |
| I use the SET and a self-designed form with open-ended items for each course I teach. I want information about teaching from any class I teach. The SET gives me an overview and a set of numbers I can give to my chair. But it doesn't give me the depth that I want to know about what I'm doing--not only what I'm doing wrong, but what I'm doing right too. |
| I find my self-designed forms really valuable. To have the students talk about what they see as important in the course and what they don't |
Process of Use

Regardless of format used, the faculty generally reported conducting written student evaluations at the end of their courses. Only two faculty members said that they had ever administered a mid-course student evaluation, and both of these used a descriptive format.

The users of rating items tended to administer the same items for every course, every quarter, and every year. The users of only descriptive items, or a combination of formats which emphasize descriptive items: 1) indicated giving more thought and attention to individualizing the item content to the course being evaluated; 2) described more personal involvement in designing the items; and 3) expressed a much stronger personal commitment to using student evaluations of their teaching and courses as an important tool for self and course development and improvement.

The users of written student evaluations reported that, after administering the evaluations, they: 1) review rating results for significant trends; 2) review any written comments and consider possible changes to make; 3) seldom discuss the results with anyone; 4) often provide a summary of the results to the chair, although the chair does not usually discuss it with them; and 5) file the ratings data for annual reviews and/or future promotion and tenure proceedings.

The judgment phase of the written student evaluation process appears to be similar regardless of format used. The faculty reported that they, for the most part, individually review the results and individually decide whether or not to make instructional adaptations/improvements. There were only six faculty members who reported that they had ever included
someone else in this decision process. These were all faculty who had either team taught a course or had a TA or lab demonstrator.

Users of rating items said that they simply look for items which a "large number" of students may have rated below average. The faculty weigh the students' responses against their own internal values and decide whether to make related changes.

I guess I probably look at the mean rating. I'm merely trying to pick up those things they rate less well. I have to then ask if it was one of my objectives to do that. (f02)

They have pointed out, confirmed problems with the course that I know exist, but they are not things that I am willing to change about the course. (f04)

It's easiest to make changes when you feel that way yourself. (f06)

Frequency of a particular response and held values also influence the judgment process related to descriptive items or written comments. In addition, the nature of this information allows at least three additional factors to influence the faculty members' judgment process. Faculty report that they can often tell which students were responsible for particular comments. Therefore, their perception of the student influences the value they place on the comment. Second, only one student may make a particular comment, but if it is considered to be particularly interesting or significant, the faculty member may decide to make a change. Third, written comments tend to be very specific and/or explanatory so faculty believe the information provides them with more direction for making changes than do rating items.
**Reported changes.** The majority of users of written student evaluations, even when they criticized the value or usefulness of written student evaluations, claimed to have at one time or another made changes in their courses related to student evaluation of teaching results.

The changes were often described by the faculty as relatively "minor." Nonetheless they often described a process whereby, based at least in part on written student evaluations, they had changed some aspect(s) of their instruction—tests, delivery, content covered, organization of presentations, etc.

Seven of the 33 users did not claim to have made changes based on written student evaluations. These faculty were three Cornell users and four SET users.

The following are samples of faculty comments about changes made as a result of written student evaluations:

I was consistently getting feedback about a certain amount of unfairness in examinations. Since then I've overcome student objections to exams. I used the college teaching consultant to help me. (f02)

The area from which I got the lowest score was arousing student interest. Number 15 says: "Does he put his material across in an interesting way?" So I tried to address that. (f08)

Everyone jumped on me and said it was too much work for a four-hour course. They said the one thing that would really help was a one-hour recitation preceding the lab. So now we do that. (f25)

As a result of student feedback, I've made mostly curricular changes. Based on student comments I have revised the syllabi in terms of giving extra time to a subject, or I have supplemented that with programmed units or other ways to
Interact with the curriculum within microcomputers or with learning activity packages or we have a whole series of mastery tests that are nongraded activities and so we have given people practice exercises with which to practice their skills. (f26)

I get some good comments. Constructive criticism. Not that you are wonderful, fantastic. But slow down, do more overheads. Give us a handout. You need to structure this into two lectures instead of one. We would like to see a representative from this area make a presentation. Those are the things I look at. (f34)

Yes, I've definitely changed things as a result of both formal and informal student evaluation processes. One of the things that was marked a little lower was transition from one subject to another and flow of the courses. So I've made some changes this year that revamped those lectures. (f35)

You don't change things from student evaluations as much as you'd like to think you use them. I have changed some things as a result of student input, though. In one class we're not going to have a midterm, because it's a three-hour course and it meets three times per week and they said they're always behind. We decided to have weekly or biweekly quizzes. (f41)

I've definitely made changes as a result of comments on my self-designed form. They didn't like the way I used transparencies so I changed that. I've used some different teaching methods. They had concern about the monotony—everyday it was lecture. I tend to be fairly dry. So for instance, today I broke the students up into small groups and they discussed some information I had given them. (f03)

My lab assistant and I go over the student evaluations and try to make changes in the course. We've changed course content or emphasis of content. We used to spend a lot of time on some subject matter that the students thought wasn't really that much value to them. So now we assign that as a part of a thing that they have to do on their own. (f21)
Last year, for example, I added an additional midterm. They thought there was too much information for one midterm and a final so I've added an additional midterm. And I've changed the percentage points based on student input as to how much things should be worth. (f28)

Factors Related to the Use of Written Student Evaluations

Faculty rank. As shown in Figure 1 on the next page, an instructor's rank appears to be related to both the use of written student evaluations and the type of items used. A full professor was almost as likely not to use written student evaluations as he was to use them. Eight of the 18 full professors currently use no written student evaluations, whereas all the assistant and associate professors in the sample used some means of conducting written student evaluations. However, the assistant and associate professors were more likely to use rating items than they were to use either only descriptive items or a combination of formats. The full professors, if they used written student evaluations, were more likely to use descriptive items only or a combinations of formats.
Legend:
- Full Professor
- Associate Professor
- Assistant Professor

Figure 1. Use of Written Student Evaluation Among Three Faculty Ranks
Policy. The university does not mandate the use of any one form for written student evaluations of teaching, but one of the duties of the chairman of a department or director of a school is specified in *Rules of the University Faculty* (1983) as:

To promote improvement of instruction by providing for the evaluation of each course when offered, including a written evaluation by students of the course and instructors, and periodic course review by the faculty (pp. 12-25).

The *Faculty Handbook* (1984) further specifies "opinions of students, appropriately documented and accompanied by interpretive information" (p. 2.a.15) as one type of teaching performance evidence to be considered in promotion and tenure decisions.

The faculty reported great similarities in the ways these policies are being interpreted and enacted across academic units. They described their department policies as "encouraging but not requiring" the faculty to conduct written student evaluations. Only one department actually mandated that a particular set of student evaluation items be administered in all courses.

The faculty generally reported that they believe that quantitative measures of student opinion are preferred over qualitative evidence in promotion and tenure proceedings. One result of this belief is that assistant and associate professors are inclined to restrict their collection of student feedback information to primarily ratings data, even though they described the data as inadequate for their needs related to instructional improvement and development.
Annual Performance Reviews

Annual performance reviews were mentioned by 13 faculty members in eight different departments as a process related to instructional evaluation activities. The remaining 28 faculty did not mention annual reviews apparently for three reasons: 1) not all faculty participate in annual reviews. (In some instances only junior faculty are required to have annual reviews); 2) so little time is spent discussing instructional activities during annual reviews that not all faculty think of them as instructional evaluation activities; and 3) when instruction is discussed, the focus is so general that annual reviews do not serve instructional improvement aims.

Those faculty who mentioned annual performance reviews described them as a process through which the faculty member meets annually with the department chair to discuss past accomplishments and future goals. A written report is usually prepared by the faculty member, given to the chair, and used as a reference for discussion. A section for listing and describing instruction-related activities is usually a part of each of these reports, and written student evaluations of teaching, or summaries of the results, are most often included. The faculty were probed as to: What portion of the discussion focused on instruction? What about instruction was actually discussed during annual reviews? Did they receive feedback useful for making changes? Although annual reviews were the third most frequently named source of formal instructional evaluation, when probed, faculty claimed that very little time was spent discussing instruction in annual review meetings.
When instruction is discussed it's reportedly done in very general terms. The chair might say: "It looks like you're doing O.K." or "Your student ratings look fine--now what about your research activities?"

In addition, although faculty reported usually including their student evaluations of teaching as part of their annual reports, chairs reportedly seldom discuss them.

Following are faculty comments which pertain to the annual review process as a tool for making instructional improvements.

In our annual review we cover all aspects of our appointment--teaching, research, and service. I guess that by far the most effort is put into research. On my annual report the feedback came back that I was to focus my research efforts more. I was too fragmented. There were two statements that were kind of in all caps in the memo I received. It said: Focus your research and strive to publish the results. I found out that all four assistant professors had received the same statements. That kind of took away the meaning of it. (f03)

I guess I basically see the annual review form as bookkeeping. I don't like all the paperwork and the time it takes to accumulate all that information. But if that is the way things are done...I keep a file and I fill that out so it is bookkeeping. It enables me whenever the next promotion and evaluation time comes that information is there and it can be pulled out and summarized. It is in a convenient form. (f04)

The focus on the role as a teacher is basically on the courses that were taught. You taught these five courses....this is the enrollment. That's about it. Most of it is focused on what you are doing in research. (f20)

Q: What does happen at your annual review conferences?
A: Not a whole lot. In our department I guess I would characterize it as going in to talk about the weather. (f22)

It is an exercise in paperwork. It never goes anywhere. (f34)
We only meet for half an hour. They can't give you much there. In the evaluation of teaching that is. As far as pouring over evaluation forms and deciding what you are doing right and what you are doing wrong, we don't do that. We are more inclined to discuss the broad scope. (f36)

Q: What about annual review time? Is teaching discussed then?
A: The most it has ever been discussed is a comment like: You are doing a good job in teaching, what can I say? Now let's talk about your research. Teaching has not been discussed very much. (f44)

I never felt like the chairman talked to me a great deal on teaching—particularly on the quality and even on the subject matter. (f39)
Analysis of Student Performance on Tests

Ten of the 41 faculty mentioned that they sometimes reviewed the test performance of their students in order to assess the quality and appropriateness of their exams. The most common form of analysis was described as checking to see whether a large percentage of their students were missing a particular item and if so they would then consider changing the item. This analysis process was usually done by simply perusing exam results rather than through any method of statistical item analysis. A few faculty reported having used the university's computerized item analysis service for multiple choice items; however, they weren't always certain they were interpreting the statistical printout correctly. The majority of faculty (76%) did not mention analysis of student exam performance as an instructional evaluation method. The following are comments representative of those faculty who mentioned analysis of exam results as a means of evaluating their instruction.

Exam time for me is an evaluation of whether I've gotten it across to them or not. If I have an exam item that only 5% of the class gets, then I say: Hey, I screwed up. Somewhere I didn't get the message across to them. (f20)

The general exam is a good time to evaluate....We don't have a formal structure for doing that, but what we would see is that if we have a failure rate that goes up then we would raise the question of why aren't our students performing on the general exam like they should? (f15)

You ask yourself why did so many students miss this? The answer tells you if you should take five minutes the next lecture to go over this. (f26)
I might say to the students: You did really poorly on this quiz. Why? I want to know why. What is the problem? What is the remedy? (f27)

We've had one midterm this quarter. That generates quite a lot of feedback. I did use the university's grading service...their exam analysis and all of that...I came out about 15 points too difficult. That was not too bad, but I was kind of concerned and there were several questions that I asked on the exam that I had to redo. I can't use them. So I've seen what those were and why those occurred. I only used two pieces of information from the printout from the testing office. All the other stuff on it was for someone else. I don't read it. It has no meaning for me. Perhaps if would if it were explained properly. (f31)
Departmental Interviews/Surveys of Graduating Seniors

Ten faculty, representing six out of the 11 departments, referred to interviews and/or surveys of graduating seniors as an instructional evaluation tool which was either operating currently or had occurred in the past. The faculty reported that the chair conducted these surveys for the purpose of assessing the students' opinions about individual courses, faculty, and their program of study. Even so, in many instances the information was collected but not shared with the faculty. When it was reported to faculty, the information dealt with instruction in such general terms that the faculty were unable to use it to make changes. They did report that it gave them general impressions as to the students' level of satisfaction with a particular course or program and, when that feedback was positive, it served to reinforce or maintain some of their practices.

Although none of the participants reported receiving negative feedback through this process, they were under the impression that if the information reported to the chairs was particularly negative the chair would inform the instructor. Hence, they viewed it as a process that potentially could lead to instructional improvement. The following statements were typical of those which referred to the exit interview/survey process.

Our chair has a form for seniors. Each student completes a series of ratings. He [the chair] shared this with us. I'm sure he uses that somewhat in salaries. He's made changes in teaching assignments based on those evaluations. He's switched people out of teaching based on continued negative ratings. (f02)

For a long time our chairman asked the graduating seniors into his office and he sat down with them for two, three, or four
hours and asked for comments on courses, suggestions for improvement of courses, faculty performance, etc. Those were used for promotion and tenure I think. The chairman would just take notes. He would list course numbers and ask them what courses they had taken.

Q: Does he then share that information?
A: You will have to ask the chairman what he does with it.
Q: Do you ever get any feedback at all?
A: No. But I think I would have if there had been real problems. (f16)

For a while the chair was doing senior exit interviews. Sometimes he would share that with us. The students filled out a form and in general evaluated each course in their major. There was a spot to evaluate each professor in the unit. I don't think there was ever 100% compliance in terms of the students that participated. One student liked the course and liked you and there were other students that were sour grapes. They thought they had worked too hard so they would slam you. (f17)

Our chair makes a point of talking to each person in the last quarter before they leave the university. He tried to get some input from the students about in general the different courses he'd taken and his evaluations and suggestions of them. When we have our annual evaluations, he gives us whatever feedback he has from that source. I haven't really gotten anything that's been very much help with reference to teaching. I guess I haven't heard anything that's been extremely critical, or the information he gives me is not particularly specific. (f18)

In the past there have been senior exit interviews. But they were never used. If they were really bad, the associate chair might go to the faculty member and say: Here's something that concerns the students. But they were not taken forward year after year if the person was rated low. (f37)

We have what we call graduating senior exit interviews. Our chair samples a few students and they have exit interviews and a written instrument on each professor. When we first began these, I could see the results—sort of a synopsis of how I stacked up with the others. Now they're available as I
understand it in my boss's office, but in my faculty counseling and review sessions each year he says: Well, you know you do pretty darn well in teaching. If you want to see the exit interviews, fine. But your teaching's fine and great and you have to focus on your publications. (f42)
Discussions with Individual Faculty and Teaching Assistants for the Purpose of Critique

Six faculty mentioned conditions under which they had held one or more general discussions with another faculty member for the explicit purpose of reviewing a course. Two of the interviewed faculty once team-taught a course together and now rotate teaching the same course. They each independently acclaimed the value of the feedback gained from one another and felt that their discussions had shaped and reshaped the course several times. They both said they wished they had the opportunity to have the same kind of interchanges about the other courses they teach.

I team taught a course once and we would discuss it. I've also had discussions with my lab TAs. The most useful source was probably my interactions with the faculty member who also teaches the course. We also talk about his courses that are complementary to mine. (f06)

About every two years I sit through the section of the course that's taught by the person who I team taught it with and who now teaches it once a year. I try to see how he's doing things and we talk about the course. He sits in some on my section too. I'd like to have people come in and see what I'm doing. I'd like the feedback. (f24)

One of the these team teachers also described a situation in which a faculty member from another department who was interested in seeing what content was being covered had sat in on one of his courses for the whole quarter. He said that they had "a very useful discussion" at the end of the course.

I taught a course once and a person from another department came over to be sure there wasn't any material that was in
conflict with what he was teaching. We had a good discussion afterwards. He sat through the whole course. It was the first time it was taught. (f24)

Four faculty members said they discussed aspects of their course with their teaching assistants or lab demonstrator. They described a process through which, at the end of the quarter, they sat down together and talked about how things went and what they might want to change next time.

My lab demonstrator and I go over the student evaluations and try to make changes in the course. We've discussed, along with another person, what we think the students need to know and what we need to do to teach those things. I share my tests with another person who teaches the course with me [the lab demonstrator] and we look over each other's tests to see if we think it's a fair test. (f21)

Immediately when the course is over, we [faculty member and TA] look at what lecture should be emphasized, what should be deleted. We do this immediately afterward while the feedback is still fresh. (f13)

I have discussed my course with the TA who helps teach it. We've talked about what should or shouldn't be changed. (f41)
In-Class Observations by Peers (Planned Observations Resulting in a Written Report)

Five faculty reported instances in which a colleague had observed their teaching performance and had made evaluative notes. In three of these cases the reports were given to the chair rather than the faculty member and the information was assumed to be used as a part of personnel decision-making rather than course improvement and development.

I was observed...once by a faculty member who was on a committee assigned to observe me when I was first up for promotion and tenure. He and the chair were the only two from a committee of four who actually visited my class. I guess they used that for P&T. (f06)

There was a committee set up to evaluate my teaching. One guy came to ten classes, the department chair came once...and then we never met. I guess they decided I was doing O.K. I don't know otherwise. (f24)

Professor ---- was in my class and I asked him to stay...and write up a report on how things went which he did and then he shared that with my chair for my file. (f27)

A fourth faculty member reported that he had a joint appointment in a department associated with the education discipline. He said that once or twice a year he requests a member from that department to observe one of his classes and write an evaluation. He described the process as one in which they use no particular instrument, but rather report generally about dimensions of teaching "such as clarity, enthusiasm, businesslike manner, etc." He reported that the feedback is usually positive so he doesn't need to change anything.
A fifth person, an assistant professor, reported that a senior faculty member sat in on his entire graduate course the third time he taught it, because the senior member was interested in learning the content. Toward the end of the quarter, they discussed the possibility of the observer writing an evaluation report. While the assistant professor was interested in the feedback for improvement purposes, he reported being primarily interested in receiving a written report by a senior faculty member that he could put in his file for promotion and tenure. He didn't describe having made any particular changes as a result of this observation process.

I asked a senior faculty member to sit in on my course last quarter. He was interested in the content. It's up to the individual faculty member, but when you are starting to look towards your four-year review, those are the kinds of things you need besides the Cornell. It is wise to have some full professor sit in and review your course—review both presentation and content. You don't see a lot of it happening. It is very expensive in terms of personal time. So the circumstances have to be right. (f40)

In-class observation by peers, whether strictly for promotion and tenure or for the individual development of the faculty member, was infrequently cited as an instructional evaluation device. When it was cited, the technique was reportedly used only episodically and had never included standardized or non-standardized observation instruments.
Group Discussions with Teaching Faculty for Purpose of Critique

Four faculty mentioned that they have sometimes discussed courses with other faculty in their departments, or with faculty involved in teaching interdepartmental courses. Faculty reported that these discussions sometimes included sharing informal student comments about related courses both within and outside of their departments, as well as sometimes sharing course materials. Again, these discussions were characterized as infrequent and general in their focus. The discussions were reported as sometimes occurring when faculty were trying to coordinate a series of courses across departments, when a new course series was being instituted, or when there were several faculty teaching different sections of the same course. The large majority of faculty (90%) did not report current or past involvement with colleagues in such discussions.

Sometimes we will get together in faculty groups and we will look over course objectives or outlines and talk about subject matter. Rarely will we discuss teaching method. We will discuss visuals or something like that. (f29)

I have talked with faculty who had the next course above mine. I'd say we have a fairly good dialogue going. We share our course materials on a fairly regular basis. We share the material that we teach with the teachers of the other 100 sections. (f44)

We have three people teaching each of the courses in our 700 series. I taught part of it. In trying to bring the whole thing together, we do sit down and we talk about these evaluations and we look a little bit at course content. When you start up a series there are some things that work and some things that don't work and so you have to make adjustments. So we get together as a committee or as a group--whoever is teaching in
there—and we hash those things out. The course is about four or five years old and the discussions are still going on. I would say that we are pretty close to having it where we want it now.  

Dr. ---- and I have been working with this particular course for 16 years....We have developed common learning objectives and a course handbook. We tried to give the same tests but it didn’t seem to work well. With lectures you tend to say things a little differently and that can affect the student’s interpretation of the questions....We developed this handbook. The sections that I have done Dr. ---- has reviewed and vice versa. (f36)
**Curriculum Review Committees**

All 11 departments have standing curriculum review committees. Four faculty mentioned these committees in the context of discussing instructional evaluation processes. The activities of these committees were described as reviewing course outlines and syllabi for all new courses proposed within the department. A member of one of these departmental committees gave the following description of the process:

>If there is a faculty member who wants to start a new course, he or she will prepare the documentation for it and present it to our committee. We will discuss it, look at it, see if it is something that is needed in the department that we have the resources for. We are particularly concerned with duplication—being sure the proposed content isn't already covered somewhere else. We look at prerequisites to be sure they're reasonable for the course level, credit hours versus proposed meeting hours, that type of thing. If we agree it's needed, we'll forward the paperwork to the college's academic affairs committee and they'll do a similar review. (f11)

If the proposed course is turned down, the proposal is reportedly returned to the faculty member for a revision of the documentation. These committees allegedly function for the primary purpose of screening new course proposals before they are sent to the college's academic affairs committee. There was no evidence from the interviews that indicated that once a course was accepted it was ever again reviewed by the curriculum committee.

Additional departmental committees were described as sometimes conducting curriculum reviews of established sequences of courses or programs of study. These reviews were described as irregular (episodic) in
occurrence. In other words, these department-initiated reviews were not
described as occurring on any scheduled basis, but rather as a response to
some perceived situation or crisis. Faculty in three departments described
such reviews as being conducted during the academic quarter in which the
interviews occurred. None of the faculty in any of the other eight
departments referred to any such process occurring either currently or in
the past. In two of the cases where the reviews were being conducted, they
were described as stimulated by a college review of graduation
requirements, conducted by the college academic affairs committee and
initiated 12-18 months prior to the interviews. One interviewee, who
serves as a member of the college's academic affairs committee, said it
was the first such curriculum review in the 11 years he had been with the
university. Faculty described these departmental curriculum reviews in the
following terms:

I think there were some rumblings among faculty that we
weren't looking at the programs regularly, that technology was
changing and we had not kept up with it and our students were
not as well prepared as other students. Something that added
impetus to this was that we got a new chairman in charge of
academic affairs. It was a way for him to get a good review
and to establish has ground rules. What we're trying to do is
put more rigor into the area by coordinating things better so we
don't have duplication, etc. Most of it was based on: Are we
giving this too much repetition in course after course? Do our
students really need a basic understanding of this principle or
do they need to know how to apply it? (f37)

A faculty member from another department said:

We're saying let's pretend we don't have any courses on the
books and where would you start? We're looking at sections,
sizes of classes, etc. I suppose from the beginning of time it's
never been reviewed. It's mostly been add courses and nothing's ever been dropped. It's kind of piece-meal. I would say the program won't change dramatically, but in some cases the people teaching the advanced courses don't even know what's going on in the intermediate courses, etc. We only have faculty meetings about once a quarter and when we do we don't really discuss the curriculum. We don't ever have a seminar or retreat on teaching. The review process will give us some reason to discuss what's going on. (f41)
Discussions with Individual Students or Groups of Students for the Purpose of Course Critique

One faculty member said that he had, so far, taught only one offering of a formal course. At the end of the course, he had asked a small number of graduate students to meet with him and discuss the strengths and weaknesses of the course. He characterized the process as personally useful, but did not cite any actual changes made as a result. Another participant said he used to have a group discussion with his students at the end of the quarter, but he stopped doing that because he felt the course was well enough structured now that he didn't need to.

I used to at the end of the quarter have a session where I'd ask them to tell me what should I be doing differently. I don't do that much anymore. After 11 years I'm at the point where the change is more gradual now....A course does get developed. And if we are pleased with it, we stop hassling people about it. (f27)
Individually Requested Review of Materials

Only four faculty specifically mentioned that they had ever requested their course materials be reviewed by anyone else. Two of the participants produced a lab notebook and asked either their chair or a colleague to review it. Another participant said he recently turned all of his course notes over to a professor who teaches the first course in the same series of courses. He remarked that: "He's already done some things to improve the organization of the material." (f10) The fourth participant reported that he and another faculty member had developed a course manual together, and they had reviewed each other's sections.

In the case of by lab notebook, I've asked people to look it over---people who are knowledgeable in terms of the type of people that I have in class and in terms of the subject matter. I've asked for their comments. (f04)

It strikes me that you ought to have more input than just one person's into a course. So this quarter I handed all my course notes over to another fellow and he's going through them and hopefully after that then we can improve the course. He's already done some things to improve the organization of the material. (f10)

I have written a lab study guide--the only one of its kind and I'm writing a text that will come out next year. I gave the manual to my chair and he has looked at it very critically. (f25)

Dr. ---- and I have been working with this particular course for 16 years....We developed this handbook. The sections that I have done Dr. ---- has reviewed and vice versa. (f36)
Consultation with a Teaching Improvement Consultant

The college chosen for the case study had a faculty member who was available part-time to function as a teaching improvement consultant. Two faculty members in the sample had voluntarily used the service. One was an older faculty member who several years earlier had requested help with his exam practices. He reported that he was very satisfied with the assistance he received from the consultant and that he had changed his learning activities and exams to more closely match his learning objectives as a result of the consultation.

The other person who used the consultant service was an untenured assistant professor who described himself as wanting to be a "master teacher." During the 1983-84 academic year he had requested from the teaching consultant an in-depth review of his in-class teaching performance. He reported that the consultant videotaped one of his classes and told the faculty member that he was "doing fine." The faculty member said that he was not particularly satisfied with the consulting experience because he was "left with sort of a cold feeling." It was concluded that a compatible personality match between teaching consultant and faculty member is perhaps an extremely important factor in the potential effectiveness of the improvement consultation.
Faculty Development Committee

One department represented in the sample was reported to have voluntary "Faculty Development Committees." These were described by two participants from the department as committees which, upon request of a faculty member, are appointed by the chair "to help people improve their teaching." (f27) When functioning, they arrange "to look at our syllabi, some in and critique our teaching, talk to our students, and give us feedback." (f27) No one is required to use such a committee, but they are suggested as a good idea for junior faculty. Neither of the faculty who mentioned these committees reported having a development committee looking at their instructional practices in the academic year of the interviews. One of the two reported that he had used such a committee in the past, and both participants described the committees as useful for improving instructional practices.
**Colleague Rating Form**

One associate professor said that his department has a colleague rating form that is used once a year to rate the assistant and associate professors who might be considered for promotion and/or tenure. Although this instructor named this form as an instructional evaluation tool, the form was described as not exclusively or very directly related to specific instruction-related activities. The faculty member reported that as a member of the committee he had seen his own form and found the feedback to be "very valuable." He described the items as being "a bit heavy on the personal, but overall you can see where your strengths and weaknesses are. It has things like professional conduct, ability to get along with colleagues, research abilities, teaching abilities, and things like that." (f30) This was the first time this instructor had seen his own rating form, and he recommended it as a potentially useful tool for providing feedback to faculty.
Informal Instructional Evaluation

Informal instructional evaluation refers to the ways in which faculty acquire information about their own and others' instructional practices through the natural course of teaching and or interacting with students and other faculty and then use that information in some judgment making process.

In addition to the formal instructional evaluation methods described in Section A of this chapter, when the faculty participants were asked to describe their instructional evaluation activities they invariably mentioned a variety of informal situations which they considered to be a form of instructional evaluation. In fact, several faculty stated that they considered the informal process(es) to be more dominant, and in many cases more meaningful and influential, than many of the more formal instructional evaluation activities in which they had participated. The participants described the informal information acquisition process as contributing to their evaluations of their own instructional practices, their personal assessments of other faculty, and their chair's evaluations of faculty performance.

These informal events, as described by the faculty, most typically include: non-verbal communication between the faculty member and the student(s) in class; comments made by students during classtime; conversations with students in the faculty member's office, at social events or at conventions; conversations with other faculty or chairs.
The information received through these situations was never formally documented, but the faculty explained that they considered these situations to be a form of instructional evaluation because the information acquired through these channels is influential in both improvement and personnel decision-making.

**Self-Assessment.** In terms of self-assessment, the faculty seemed to consider the informal, or natural feedback process, as an important day-to-day evaluation tool. They reported that they consider what happens in the classroom as an important barometer for judging and adjusting their instructional behavior. They explained that they use the students' facial expressions and daily interactions to help them decide if it's necessary to change or maintain their course delivery and organization.

I'd say I evaluate my course day to day. Because I think the best evaluation is sitting there staring eyeball to eyeball and saying are you getting this or are you not getting this? I'd say that I can pretty well look at a class and see whether this is mystery hour or they're understanding what I'm talking about. I give them lectures...on how important it is to ask questions if they don't understand something. I really encourage questions—that and with the looks on their faces tell me pretty well where I am and where they are. (f01)

You listen to what the people in the class say. You give them the atmosphere to be able to say things and then you listen to them. (f05)

You've got to watch their faces. I can see if the eyes glaze over I know that I've gone past them....I call this the windowshade effect. There is a wall between the student and me. I've lost him and I've got to get him back. (f16)
I guess I found if you keep your eyes open and ears open the students soon tell you whether something you are doing is coming across to them and whether there are other ways to do it. I try new things over a time and change things. (f19)

I get kind of a feel for the class. I've been at this long enough that if things are clicking along the students and I get along pretty well. You can just feel it in class. When it is not there you don't feel it in class and so from stuff like that I can pick up on sore spots and it's the sore spots that I then ask about on the written student evaluations. (f22)

I came here with the intent of starting a good research program and getting it going. But my teaching qualities had to keep pace simply because you get immediate feedback. You screw up in a lecture, they are on you the next day. Instant feedback. You hear: Lousy lecture. That guy didn't know what he was talking about. Disorganized. I remember as an undergraduate I was one of the first ones to complain if the person wasn't doing a good job teaching. So I have this moral responsibility and I don't want to be a hypocrite so I've got to keep my teaching act in order. (f380)

I don't have to wait until the end of the quarter to know....If they have concerns with the readings, when they are confused and want to be helped, then I will know there are problems with certain students or certain readings. (f12)

I try to get daily feedback by analyzing or comparing how my presentation went—lecture-wise or discussion-wise—compared with the verbal and non-verbal feedback I get from the students. (f34)

The faculty also mentioned conversations with students outside of the classroom as a source of information about their instructional practices. In fact, some faculty reported that they consider these informal interactions as particularly valuable feedback.
I rely much more on informal discussions with students which I have frequently throughout the quarter. If they are going to prepare a term paper or lead a discussion, I encourage them to have a conference in their process of preparing it. Most of them do. (f15)

My door is always open, and students do come in frequently to talk with me about the course. I really think that's the only way to teach a course....I have 15 or 20 students and I can do that. I can get acquainted with each one of them and I can find out what they're feeling and how they're getting along. (f14)

Sometimes if they come in here to talk about some questions on their courses, I'll ask questions that will give me a lead into what they are thinking about. Based on that we can make some changes also. (f22)

The real evaluation comes, not in what they write, but in what they say when we have a couple of beers and let our hair down and talk to them after they've had the course and find out what they really think. Sometimes I'm amazed. If I hadn't taught the course I'd swear they never had the course. It becomes very much a discussion with the students. (f24)

If a student comes in for counseling or to sign their schedule I will ask them: "How was the class? What did you think of the lecture today? Was what I said last week clear or should I go back and repeat it?" I select those students carefully. I select the students that are more willing to talk and are willing and aware of getting knowledge and are not grade conscious. (f34)

Students will come in and talk to me, and I'll ask students. I'll be walking down the hall with them and I'll ask are you beginning to see what I'm driving at? (f42)

I develop a personal relationship with all my students, and they will come in and talk with me. So I'll ask them for their comments. (f43)
I like to know my students will enough that I can call them by their first name. They're usually frank enough that if they have an idea, they'll tell me. So they can make suggestions anytime. So I don't know that I learn a great deal from these evaluation forms. (f45)

I know the students pretty well. We're on pretty good terms and you get a pretty good feel for what's going on. (f41)

A limited number of faculty described informal interactions with colleagues as a means through which they have received information about their instruction. While this information was reportedly transmitted through colleagues, it usually consisted of comments which had been made originally by students. In addition, the information gained through informal interactions with colleagues appears to be very general in nature and in the interviews was never related to specific improvements or changes.

Occasionally a faculty member may mention some comment he's heard from one of my students—whether they thought things were or were not going well. You get a little bit of that. You could probably get more of that if you really solicited it. I haven't done much of that. (f08)

Feedback comes most directly from students, but eventually it is fed through colleagues. I mean eventually colleagues would feed back the same thing that I had already gotten from students. (f19)

I've talked to people who had had graduate students in my course and we've discussed a little bit about what the students have picked up. That's pretty informal though. They've been pleased with what they've gotten and it's always nice to hear. But if hasn't had much impact on what's been offered in the course. (f24)

I know my students talk about what I do in class. They will discuss that with their advisors, and sometimes the advisor
will come to me and say: "I guess you're doing a good job in this class, because you've got somebody on their toes." (f34)

Colleague Assessment. While informal interactions with colleagues do not appear to contribute directly a great deal to the faculty's self-instructional improvement and development effort, informal interactions with both students and colleagues do appear to influence the faculty's opinions of their colleagues' instructional practices. The faculty frequently reported that they form impressions of other faculty through what students tell them and through their contact with their colleagues. They reported that they seldom approach other faculty regarding negative impressions but that they might tell their chairs, or they assume the chairs hear the same things and that the chairs will address any instructional problems that emerge.

With regard to undergraduates, if there is any involvement in terms of social gatherings or visits to the professional club, there are always comments made about courses and instructors. I think that administrators through the grapevine maybe do find out something about the quality of the job that various faculty members are doing with regard to instruction. (f04)

There's lots of folklore about who's a good teacher and what's a good course. Students make comments to other faculty or to each other in the hall and that influences people's perceptions. It's very unscientific, but I know I'm guilty of being influenced by it. (f06)

The chair gets the news from students or he gets the news from colleagues in the department who have heard from students if something is wrong. I think maybe the really bad cases are taken care of. (f10)
We have social events and professors invite students over to their houses. In a social setting such as this if a professor was screwing up it wouldn't take long to get the chairman's attention in this department. There are several assistant professors in our department that I have heard a whole slew of comments about. Some of the students are my advisees and they don't hesitate about giving me a long lecture about various professors....When a professor is a poor professor I can tell that when he is giving a public talk--he is not articulate, doesn't have a command of the English language, will frequently engage in grammatical errors that are not acceptable, his command of the material is questionable--I find it a bit difficult to go in and tell him this. I would not hesitate to discuss it with him if he brought it to my attention. (f12)

When I have seven or eight students in here to complain about a class, the lecture, the test questions, the test...you have a real problem. I have gone to the chairman on a case like this. I wanted to make sure my students wouldn't waste their time on taking this course from him. In this case the students are influencing the status of a professor from word of mouth to their advisors. The feedback is valuable. The students provide more of an immediate role in influencing faculty attitudes by going back to their respective advisors. If something is really rotten in Denmark it gets out pretty quickly. (f12)

Where there are problems they surface. The problem instructors have been called into the chairman's office who describes the way it is. We hear about these things of course. But you are very careful about poking your nose into troubles of peers. That's not your job.

Q: How do these problems surface?
A: From students. Where it ought to be. Our graduating class is 15-18 students and you get to know them very well....When you have a small group like that you get cohesive. If they get upset with an instructor this cohesive group will corner the chairman in his office and let him know...which is the way it ought to work. (f16)

By the time you've been here five or six years, everybody knows about your teaching. I don't think there's any hidden
Information about how good you are. Intuitively, we all have a feel for how good a faculty member is because we hear even though we never step in the classroom. I think 80% of the judgment about your teaching is made from just "knowing" you're good or based on what students have informally shared around the department. (f25)

I have students that sit in that chair and cry their eyes out over exams that weren't fair. "I took that course and couldn't believe that exam. I knew the material. I studied 25 hours for that test. I knew the content and every question was ambiguous on that exam. Sixty percent of the students failed and the professor is going around in the corridors bragging about it." I get irate. That's not good teaching. That's poor teaching.

Q: Did you ever intervene in a situation like that?
A: Personally I do. Professionally I don't. I have looked people up and told them I didn't think that was fair. They paid no attention to my comments I'm sure. Professionally I don't interfere. (f26)

If there's a good course on campus you will hear it in this office probably 15 times a day, and if there's a bad course you'll probably hear it 50 times a day. That information eventually filters back to the instructor. The informal evaluations are probably the most powerful. (f28)

You learn a lot about other faculty on a five-day field trip. They [the students] get to know you and they sit beside you all the way to Chicago. Pretty soon they're pretty free with their comments. (f37)

If we have an instructor that does an exceedingly poor job, we just don't send students to him. We learn about it from our students. (f39)

If evaluation is too planned it gets too formal. If you are having a cup of coffee and something comes up and you happen to mention it—sometimes you get more information that way than you do otherwise. I learn a lot that way. (f39)
With our contact with our graduate students we know pretty well what is going on in other courses....I wouldn't be surprised if through graduate students we don't know more of what is going on, because we are not sitting in on the course itself. (f39)

I know most of the other teachers in the college. I think that when you know a teacher and talk to him about his course you know if he is doing a decent job or not. (f44)

You soon find out about other faculty by talking to the students. Some professors ask each student how they got along in certain courses. I do a little bit of that if I know students are saying things like I'm wasting my time in so-and-so's class, or I should never have taken that course, etc. Then I might ask some other students. (f45)
CHAPTER 5: INTERPRETATIONS

Introduction

This chapter is intended to offer an interpretation and critique of the results of these interviews. As such, it reflects the values and experiences of the particular author. While alternative interpretations could probably be made of the same descriptive base, it is intended that the interpretations offered are justified by the findings reported in Chapter 4 and the arguments provided in the following sections.

Faculty as Instructional Decision-Makers

The comparisons of universities to "organized anarchies" (Cohen, March & Olsen, 1972) and "loosely coupled systems" (Weick, 1976) seem particularly appropriate as they relate to the realm of instructional decision making. According to these interviews, it is primarily individual faculty who are deciding what subject matter is taught, how to teach it, how to evaluate student performance, how and when to evaluate courses, and if and when changes or improvements should be made. The word "individual" is emphasized because, in most cases, it is individual faculty--alone, without the input of others--who are making instructional decisions.

This individual decision-making appears to begin with course design and to continue as the course is implemented and evaluated. For example, with
the exception of one department which required that a uniform group of S.E.T. items be administered in every course, it was essentially up to the individual faculty member to decide if and how information about instructional practices would be collected for both personnel decision-making and improvement and development purposes. Occasionally departmental faculty will convene as a group, or in smaller working groups, to review a department's curricula or sequences of courses. Even so, this type of discussion apparently occurs so infrequently that courses usually operate for years with no functional review except by students. In addition, departmental reviews appear to focus on very general content structure across a department's offerings rather than on issues internal to given courses. Many of the most critical instructional decisions are being made by individual faculty and include others only if the faculty member seeks the involvement of others.

The processes of instructional decision-making as described by these faculty, therefore, confirm the rationale that in order to understand how instructional evaluation directed at improvement and development operates within a multiversity, it must be examined from the perspective of individual faculty members.

**Instructional Evaluation as a Concept**

When faculty participants were asked to describe their involvement in "instructional evaluation" activities, most of them initially seemed automatically to associate the term with written student evaluations of teaching and personnel decision-making. To professional evaluators and instructional researchers, the term instructional evaluation stands for a
complex set of inquiry methods aimed at multiple decision-making purposes. In contrast, most university faculty have not been exposed to the literature on instructional evaluation, and they hold a more simplified concept of instructional evaluation which reflects their personal experiences in their practical worlds. They apparently learn to view instructional evaluation as primarily the process(es) of collecting student ratings of teaching data for the purpose of "documenting" instructional performance for promotion and tenure reviews. This acquired concept of instructional evaluation reflects several realities of the faculty member's organizational life.

Unless their academic discipline is education, most university faculty have not formally studied instructional design, delivery, or evaluation. They have been busy studying and staying current in their own academic disciplines, and what they have learned about planning, implementing and evaluating instruction, they have learned on the job, or from observing their own teachers through the years. Most faculty, therefore, enter their teaching role with a relatively undeveloped notion of what instructional evaluation is or could be. Some faculty most likely begin their faculty positions without having ever thought about or used the term "instructional evaluation." For others, completing ratings instruments for their own professors, or conducting student evaluations as TAs, may have been their only exposure to the idea and practice of evaluating instruction.

Within the climate of today's university, this narrow concept of instructional evaluation is either quickly learned by or reinforced in the new faculty. The following description is an imaginative scenario of the process through which a new faculty member may learn and internalize the concept of "instructional evaluation."
**Scenario**

Ms. Peacock was a lab assistant during her graduate studies and has been hired as an assistant professor at Greenley University. Upon arriving at the university, Ms. Peacock is told that her performance as a teacher will be considered in future promotion and tenure reviews and that she therefore needs some way of "documenting" her teaching. She is given no strict rules for this documentation process, but her chair says it is a good idea to conduct some sort of a student evaluation of each of her courses and to save the results for future performance reviews. She hears via the "grapevine" that student evaluations of teaching are used in promotion and tenure proceedings, and she notices in the Faculty Handbook that "opinions of students, appropriately documented and accompanied by interpretive information" are acknowledged as an acceptable type of teaching performance evidence to be considered in promotion and tenure decisions.

Ms. Peacock asks another assistant professor what he does to evaluate his courses. He tells her that when he first came here he used to ask his students a number of open-ended questions specific to his course. He thought the information the students provided him was useful for trying to improve and develop his course, but the chair told him that some sort of quantified data was preferred for the promotion and tenure process. He explained that he, therefore, quit using his own questions and had borrowed some rating items from another faculty member. He also
explained that the university offers a student evaluation of teaching service through which one can select rating items and have them machine scored. He said they return a summary of the statistics to you and then you can file those for P&T.

The new faculty member asked if she could make a copy of his rating items to use for evaluating her course. He agreed, and Ms. Peacock proceeded to administer the items at the end of her course, to look at the results for any seriously low ratings, and to file the results for future promotion and tenure reviews. Thus was Ms. Peacock’s concept of “instructional evaluation” first conceived.

This concept of instructional evaluation was further nurtured and developed throughout Ms. Peacock’s first four years at the university. Between the time she came to the university and the time of her first P&T review, no one other than students had ever been involved in critiquing or reviewing her instructional practices. In fact, aside from a casual conversation now and then, Ms. Peacock had never discussed teaching or her courses in any detail with her chair or colleagues. Methods of instructional evaluation and their relationship to course development and improvement had certainly never been discussed in her department. When the term evaluation was mentioned by the chair, colleagues, or central administrators, it was usually linked to the word “faculty” or “performance” rather than instruction. So Ms. Peacock had naturally tended to think of evaluation as
something that was done to her in order to assess how well she had performed.

Since she was allowed to select the student rating items that would be used to evaluate her performance, she was inclined to select items that would reflect her strengths rather than her weaknesses. In fact, on her first evaluation the students tended to rate her examinations as unfair. So she eliminated items pertaining to examinations lest she get low ratings. One time she had about three students who gave her rather poor ratings overall. She thought she knew who the three students were and that they were simply lazy learners who were angry because they had to work so hard in her class. Rather than have to explain all that to some future tenure review committee, she decided to simply remove the three forms from the stack before the results were summarized.

Now, Ms. Peacock's decisions to remove the items and forms may sound a bit questionable, but can one really blame her? After all, she was under the impression that her worth as a teacher was going to be judged based on those evaluation results.

Ms. Peacock has a friend, Mr. Adams, in another department. His department mandates the use of a uniform group of rating items, and the forms are administered and collected by a TA at the end of every course. Therefore he cannot remove items or manipulate the results. Just the same, because student ratings of teaching is the only official departmental instructional evaluation procedure and the results are used in salary and P&T
proceedings, he too views formal instructional evaluation as a process that is linked primarily to student evaluations of teaching and personnel decision-making.

This scenario represents situations that were either described or implied by a number of the faculty interviewed in this study. It is intended to illustrate some of the conditions that have led many faculty to view instructional evaluation as primarily a process whereby students rate their teaching performance for the purpose of promotion and tenure.

**Problems with Current Use of Student Ratings.**

The faculty's accounts of their practices and viewpoints regarding student ratings of teaching led to the conclusion that the institutionalized use of student ratings needs serious review.

**Historical perspective of student ratings.** The practice of systematically and formally collecting student opinions of teachers and courses swept through this university and many others in the late sixties and early seventies. It emerged as a response to student outcries and demands for relevancy, humanism, and shared authority in their higher education experience. The students' demands fed into the institution's external political pressure for accountability and the management scientists' tendency to address organizational problems through systematization, centralization and quantification. Waiting in the wings were the educational measurement specialists with their armory of measurement expertise and instruments.

It is understandable why, at the time, systematic student ratings of instruction seemed an attractive solution to a complex set of issues. Ratings instruments allowed students' voices to be heard in a quiet and
controllable manner, institution-wide standardized data could be collected about the teaching enterprise, and student opinions could be reduced to simplified and manageable numbers.

Even as student rating systems were being proposed, some university faculty were not easily convinced of their merits. An historical review of OSU documents from the late 1960s and early 1970s revealed several references to faculty concerns regarding the potential organizational misuses of student ratings data. (Minutes and correspondence from Council on Academic Affairs Committee, 1968-1984). Several faculty felt strongly that the primary goal of any student evaluation process should be instructional improvement, not personnel decision-making. They expressed doubts that student opinions reduced to ratings could adequately inform faculty about needed changes, and they feared that the simplified numbers might eventually be used by administrators for P&T purposes (CAA 1968-1984).

Judging from these historical documents, it appeared that the OSU administration was listening and understanding their faculty’s concerns. The Provost initially stated that instructional improvement would be the primary objective for instituting a student rating system (1968). He agreed that use of the student ratings system should be voluntary and that individual ratings results would be provided to the individual faculty member only—not administrators. Department heads, college deans and central administrators would be provided group data only—never would they be given results according to individual courses or faculty.

This plan was enacted as formal university policy in 1969. The catch seems to be that within one year of the initiation of the university-wide
student rating system, it also became formal policy that student ratings could be submitted as evidence for teaching effectiveness in promotion and tenure dossiers. It appears that, when student ratings became deemed an acceptable measure of teaching effectiveness for P&T decision-making, the objective of instructional improvement became displaced.

It has been pointed out that, while both personnel decision-making and instructional improvement are necessary and worthy organizational goals, it is probably not possible to address each effectively through the same processes (Braskamp et al., 1984; Cammann, 1982).

Cammann (1982) explains that organizational feedback systems, such as student ratings systems, can serve two primary functions. The first is individual and organizational learning and adaptation (improvement and development) and the second is extrinsic motivation and external control (personnel decision-making). He asserts that "if a single feedback system is used to serve both sets of functions, dysfunctions occur" (p. 91). The results of the interviews conducted in this study strongly support this position and suggest that the current practices surrounding student ratings of teaching are not serving either function well.

Ratings Inadequate for Improvement purposes. Across these 41 interviews, the most frequently reported and consistently practiced instructional evaluation method was written student evaluations of teaching, with 80% (n=33) of the faculty reporting that they use some type of written student evaluation. Of those 80%, 64% (n=21) administered primarily rating items only. All of the untenured assistant professors and the great majority of associate professors reported using rating items and named promotion and tenure, not instructional improvement and
development, as their intended data collection purpose. They consistently said that they did not find the rating results particularly useful for identifying the need for, or making, course changes, but they did consider students' written comments to be helpful. They explained that they administer the ratings because their administrators seem to prefer numbers, and the ratings formats are a relatively easy and painless means of collecting such measures. One young professor expressed it this way:

If I didn't need to use a standardized form for promotion and tenure, I would do my own forms. I did that once at the same time I did the Cornell. It was much more useful information for me. The reason I quit using it was that it was too time-consuming. You have to go back and calculate all of the composite scores and then get it into a form that would be acceptable for an annual report. Then you have to explain its meaning to whoever is reading it. I think probably as pertinent as anything was the indication of the current chair at the time that the self-developed evaluation form simply didn't cut as much weight as standardized forms, because they weren't able to compare your results with the university-wide aspect of things. One of my major responsibilities is to publish to get promoted, so I don't have enough time to do more evaluation. (f24)

It appears that regardless of what administrators state in formal policy, or espouse as the improvement-directed virtues of student rating systems, the faculty do not perceive the ratings as useful for that purpose.

Ratings inadequate for P&T purposes. These interviews also call into question whether the ratings are serving well the function of personnel decision-making (extrinsic motivation and external control). Even though these faculty members aren't sure exactly how or to what degree student ratings data influence P&T decisions, they are convinced that ratings are the administrators' preferred evidence of teaching effectiveness. In this
respect the faculty perceive the ratings as linked to organizational rewards. But as Cammann (1982) explains:

If the feedback information is used as a basis for providing organizational rewards and penalties, it can increase an organization member's desire to perform well on the feedback system measures. This can provide an inducement to perform well, or to manipulate the measurement system by keeping the goals and standards low, for example, or by concentrating effort in measured areas and ignoring unmeasured ones (p. 88).

In the case of these OSU faculty, the use of student ratings in P&T decisions appears to induce the faculty to sometimes manipulate the measurement system. Not by keeping the goals and standards low, because there are no university specified standards, but through the faculty member's selection of items and their method of summarizing the data. The current feedback system, instead of functioning as a support for experimentation and learning, encourages faculty to ask a minimum number of questions about their practices and to stay away from questions that might produce negative responses. In addition, over and over again the faculty said that when the students' responses are nothing but positive the information makes them feel good and helps them during P&T reviews, but such data tells them nothing about what needs to be changed or improved. In this respect neither the faculty nor the administrators are getting the information they need.

Another weakness of the current uses of student ratings data is the lack of associated performance standards. Cammann (1982) states that any organizational feedback system "must have a set of standards or objectives that gives meaning to the information collected. These standards or objectives provide the yardstick for interpreting the information and
deciding what implications should be drawn from it...the standards must exist or the information from the feedback system will be useless" (pp. 87-88). Whether it is administrators who are trying to assess faculty performance during P&T proceedings or faculty who are trying to determine how to perform, specified standards are necessary. According to a review of institutional documents (Guidelines for P&T Reviews; Guidelines for Administering the S.E.T.; Faculty Handbook) and informal conversations with department, college, and university administrators, such standards do not exist within this university—except perhaps in the individual minds of the persons involved in the P&T reviews.

The problems associated with establishing standards from which to interpret student ratings results are not simple ones. The enormous differences among course objectives, teaching methodologies, class enrollments, student characteristics, subject matter, etc., make it very difficult to establish meaningful norms. In addition, the research on student ratings indicates that the means on student ratings items tend to be skewed toward the positive end of rating scales. This makes it difficult to make meaningful discriminations among faculty based on student ratings data. Doyle (1979) has said:

With regard to the [student ratings] statistics, a frequent temptation is to overinterpret small differences...a good rule of thumb is to disregard differences that are not at least half to three-quarters the size of the standard deviation. This will sometimes result in no differences among instructors, but to interpret smaller differences could be to create differences where there are none (p. 160)

One is left asking: If faculty find the ratings data relatively uninformative for identifying the need for and making instructional...
improvements, AND if student ratings data are so difficult to interpret accurately in personnel decision-making, why do we continue to use them, and why is their use apparently increasing on university campuses?

When one looks at the list of teaching effectiveness measures that are named as acceptable evidence in P&T proceedings at the university in which these faculty work, it becomes clear why the faculty would opt to use student ratings. Their Faculty Handbook (1982) reads:

No single set of satisfactory measures [for teaching effectiveness] can be prescribed. However, among significant types of evidence of teaching effectiveness are the following (not rank ordered): (1) opinions of other members of the candidate's department, particularly if based on class visitations, on attendance at public lectures or lectures before professional societies given by the candidate, or on the results of the candidate's teaching in courses prerequisite to those of other members of the department; (2) opinions of students, appropriately documented and accompanied by interpretative information; (3) development by the candidate of new and effective techniques of instruction and instructional materials, particularly when evidenced by acceptance at other universities; (4) publications by the candidate on the teaching of his discipline in respected journals devoted to pedagogy; and (5) recognition or awards for distinguished teaching (p. 9).

For most faculty "opinions of students" is the only measure of these five to which they have easy access. The other four measures require far more effort above and beyond their regular teaching responsibilities. In addition, student opinions in the form of student ratings are a relatively easy means of providing documentation and interpretation because they can be summarized in the form of statistics.

The university's reliance on student ratings as a measure of teacher effectiveness seems to reflect its adoption of management science's notions of efficiency and accountability. As Schon (1983) explains, most
professional organizations, including education institutions, have adopted a technical-rational paradigm for structuring and managing their activities. In such a model:

It is considered important to achieve quantitative measures of proficiency and progress which are independent of individual judgments. These are much preferred to qualitative, narrative accounts of the experience of learning or teaching. Quantitative measures permit the system of control, and the other systems that depend on it, to take on an appearance of consistency, uniformity, precision, and detachment (p. 331).

Wise (1979) has referred to this phenomenon as it operates within schools as the "hyperrationalization of education." He notes that such an approach "presupposes a rationalistic, and to teachers unrealistic, view of teaching. And it presupposes that schools function like stereotypic bureaucratically organized factories" (Wise, 1979, p. 359).

It appears that the current institutionalized use of student ratings of teaching is a prime example of the technical-rational model at work. The faculty's descriptions of their experiences with student rating instruments and data is also a prime example of the negative consequences which can result from trying to reduce important and complex human activity to a set of simple numbers.

For many assistant and associate professors, this technical-rational approach to evaluating instructional performance appears to be severely limiting, if not determining, their instructional evaluation practices. For the majority of the assistant and associate professors who participated in the study, written student evaluations of teaching have become what they describe as "a numbers game." They consider it easier to administer one evaluation format, and since numbers are preferred for promotion and
tenure, they view it as more expedient to simply use rating items and a general comments item. They use items which produce responses in the form of numbers and save the numbers to give administrators, even though they're not sure what numbers represent satisfactory performance or exactly how the numbers are used. The process becomes a game played, not to improve instruction, but rather as a means to another end—that of promotion and/or tenure. Current organizational norms appear to often result in faculty compromising the personal and organizational purpose of ongoing improvement and development for the practical requirements surrounding the promotion and tenure process.
Issues Associated with Current Evaluation Practices

Based on these interviews, it seems reasonable to conclude that, overall, these faculty have spent little time involved in formal evaluation activities conducted for the specific purpose of instructional improvement and development. The majority of the formal methods mentioned in the interviews were each named by six or fewer interviewees, and several of these were named by the same participants. In addition, many of the formal evaluation activities mentioned by these faculty had not occurred within the academic year of the interviews. They had simply occurred at some time in the faculty member's career. It therefore seems reasonable to further conclude that formal evaluation activities have not played a very meaningful role in the daily teaching lives of most of these faculty participants.

Does one, therefore, also conclude that instructional improvement and development have not occurred, or that the goals of instructional quality and instructional improvement and development are not important to these faculty? It seems not. Rather, their stories seem to suggest that:

1) The phenomenon of instructional improvement and development cannot be understood by focusing only on formal evaluation activities. The phenomenon of "informal instructional evaluation" must also be examined.

2) Faculty members, for the most part, are sincerely interested in providing quality instruction and in continuously improving their instructional practices; however, there are few organizational incentives or rewards for doing so, and organizational norms serve to inhibit such activities.
Role of Informal Evaluation

During the first several interviews, when the faculty described informal, natural events as a form of evaluation, the inclination was to discount the importance of these experiences. The initial interpretation was that perhaps the faculty were somehow trying to justify their lack of involvement in formal instructional activities by saying they at least paid attention to student reactions in and outside of class. However, over and over again, without prompting, these faculty referred to informal, natural events as the catalyst for, or basis on which, they made evaluative judgments about their courses. They also frequently stressed the positive value they place on the informal as an informative process. They repeatedly said things like:

The best evaluation is sitting there staring eyeball to eyeball and saying are you getting this. (f24)

I guess I found if you keep your eyes open and ears open the students soon tell you whether something you are doing is coming across to them and whether there are other ways to do it. I try new things over time and change things as a result. (f19)

The real evaluation comes, not in what they write, but in what they say when we have a couple of beers and let our hair down and talk to them after they've had the course and find out what they really think. Sometimes I'm amazed. If I hadn't taught the course I'd swear they never had the course. It becomes very much a discussion with the students. (f24)

I like to know my students well enough that I can call them by their first name. They're usually frank enough that if they have an idea, they'll tell me. So they can make suggestions anytime. So I don't know that I learn a great deal from these evaluation forms. (f45)
It was concluded that in the natural, practical lives of faculty, these informal events most likely serve an important role in their instructional practices. The information used for making instructional decisions appears to be often unsolicited, undocumented, and unquantified. Yet this informally acquired information apparently serves evaluative decision-making and thus effects instructional adaptations and change. It seems quite possible that what these faculty called "Informal instructional evaluation" is what Schon (1983) has referred to as "knowing-in-action" and "reflection-in-action."

The workaday life of the professional depends on tacit knowing-in-action. Every competent practitioner can recognize phenomena...for which he cannot give a reasonably accurate or complete description. In his day-to-day practice he makes innumerable judgments of quality for which he cannot state adequate criteria, and he displays skills for which he cannot state the rules and procedures. Even when he makes conscious use of research-based theories and techniques, he is dependent on tacit recognitions, judgments, and skillful performance.

On the other hand, both ordinary people and professional practitioners often think about what they are doing, sometimes even while doing it. Stimulated by surprise, they turn thought back on action and on the knowing which is implicit in action. They may ask themselves, for example, "What features do I notice when I recognize this thing? What are the criteria by which I make this judgment? What procedures am I enacting when I perform this skill? How am I framing the problem that I am trying to solve?" Usually reflection on knowing-in-action goes together with reflection with the stuff at hand. There is some puzzling, or troubling, or interesting phenomenon with which the individual is trying to deal. As he tries to make sense of it, he also reflects on the understandings which have been implicit in his action, understandings which he surfaces, criticizes, restructures, and embodies in further action (pp.49-50).
University faculty may seldom conduct or participate in formal evaluation activities which use observation instruments, survey questionnaires, needs assessment tools, or structured reviews. Nevertheless, it appears that they assess their enactment of the instructional role. Rather than suggesting that university faculty are not concerned with or involved in instructional evaluation, adaptation, or improvement, these interviews seem to confirm Schon’s (1983) theory that:

It is through tacit norms that all of us make the judgments, the qualitative appreciations of situations, on which our practical competence depends. (p. 53)....The process of reflection-in-action is central to the "art" by which practitioners sometimes deal well with situations of uncertainty, instability, uniqueness and value conflict (p.50).

These faculty's descriptions of their workday life as professionals suggests that faculty frequently evaluate what they're doing and how they're doing things but that this process is relatively unstructured, unformalized, and performed individually without the involvement of others.

Does this analysis suggest that "knowing-in-action" and unstructured, individual "reflection-in-action" is sufficient for continuously adapting, improving and changing university instruction and curricula? The answer is—probably not.
An Expanded Notion of Reflection—in Action

In the natural, physical world the image produced through the process of reflection is affected by the color, degree, intensity, and angle of the available light. The reflected image is further affected by the size, texture, shape and color of the reflective surface. Watching a child play with the image cast from the curved surface of a chrome coffee pot illustrates the role of some of these factors in producing reflected images. A house of mirrors at a carnival, a clear stream, and the experience of the woman in the soap commercial who exclaims "I can see myself in my china" are all examples of the natural, physical phenomenon known as reflection. In each of these examples, an image is projected back to the viewer, and in each case the projected image is a slightly distorted representation of the original phenomenon.

It is proposed that the ability to reflect, without distortion, on one's instructional practices is also affected by a complex set of factors. Reflection is conventionally defined as:

-a thought, idea, or opinion formed, or a remark made, after attentive consideration or contemplation; and

-consideration of some subject matter, idea or purpose often with a view to understanding or accepting it, or seeing it in its right relations (Webster, 1980).

Reflection, in this sense—as it applies in the social/psychological/organizational world—is an interplay of seeing, thinking and interpreting. These combined acts are a function of what we choose to focus on and our repertoire of beliefs, values, theories and experiences through which we make judgments and form conclusions.
Considered opinions, ideas and understandings of instructional practices can be made by individual faculty in isolation from their colleagues, students, and administrators. Still, the perspectives or "images" which result from such individualistic reflection can, at best, produce only partial understandings of the process and effects of their practices.

Instruction is a human activity that occurs in and through social interactions and organizations. It is activity grounded in a complex set of human values, politics, formal and informal theories, norms, goals and objectives. There are not rigid rules for combining these multiple factors into formulas from which faculty can derive solutions for designing, implementing, and evaluating instruction. Rather, faculty must consider each and all of these interacting factors and construct "best-choice" solutions for accomplishing particular instructional goals and objectives with particular students under particular conditions at particular times. The very essence of this complex task is human judgment. Sanders (1982) has said:

Even if humans knew enough about the requirements of the natural process of learning to cooperate with it, educating still would not be a straightforward technological task. It necessarily involves structures and interrelationships that are constructed by human beings. Human intentions, expectations, meanings and purposes are central aspects of the process (p.7).

Sanders (1982) also claims that these realities do not presuppose that instructional choices cannot, or should not, "be based on reliably grounded knowledge" (p.1).

It is, however, proposed that individual faculty reflecting alone about their individual courses is an insufficient means of acquiring reliably grounded knowledge about instructional practices within a university. In
general, the current reflection-in-action, or instructional evaluation, practices of the faculty in this study seem to be like seeking a whole body image by staring into a single 12-inch mirror tile in a dimly lit room with the shades pulled. For example, student evaluations of teaching represent but one fragmented surface from which to view a faculty member's instruction. As a process it cannot possibly provide all there is to know in order to appropriately adapt and change practices. Daily contact with students and attention to their verbal and non-verbal cues adds more area to the reflective surface, but, in order to acquire a view of instructional practices sufficient for making choices based on reliably grounded knowledge, one must have the benefit of mirrors which have been placed to project multiple perspectives. In other words, given the nature and complexity of instructional activities, collective and cooperative "reflection-in-action" would be far more desirable than the current individualistic approach.

**Collective Critique**

In order to adequately and honestly critique human actions and theories, the input or perspective of more than the actor is required. Lofland and Lofland (1984) point out that:

> Ordinary people playing out their ordinary lives are enmeshed in devising and enacting ideas and activities that are responsive to their immediate needs. They are not in a good position to assemble information, to reflect on its meaning, to envision larger contexts in which it might invariably be interpreted, or to contemplate feasible and conceivable alternatives to their situation. Additionally, the other necessary avoidances of ordinary life lead people to miss seeing many aspects of their situations, and to develop legitimizing and accommodating meanings for what they see....In other words, the requirements
of acting weaken or nullify the capacity and honesty to reflect on the actions (p. 120).

A reflective process, more adequate than is currently practiced, would include the formal "surfacing", or articulation of the theories which guide instructional actions, as well as the articulation of related values, goals and objectives. Faculty would then seek empirical evidence as "a way to determine the truthfulness of [their] claims to know" (Sanders, 1982). These theories-in-action and the supporting evidence would be discussed and critiqued with colleagues, students, and perhaps educational consultants. The theories-in-action would then be restructured and embodied in further action.

Like all professionals, university faculty bring to their organizational positions a variety of formal and informal theories which guide their practical choices. In order to function effectively, these faculty need to test the "truth-value" of their multiple theories and, if necessary, restructure them based on new evidence and insights.

All professionals could benefit from a collective, cooperative reflection process, but it is particularly important for university faculty. In other professional settings, practitioners have studied the formal theories associated with their practice. Engineers have studied the theory of thermodynamics, lawyers have studied legal and political theory, and university faculty have studied the theories associated with their academic disciplines but, in most cases, they have not studied the theories associated with education. Few faculty members have, for example, studied learning, sociological, or testing and measurement theory. They have rather derived their theories for action from their own personal experiences as longtime
students. They need opportunities to surface their tacit understandings and to expose them to external examination and critique. Schon notes that:

As a practice becomes increasingly tacit and spontaneous, the practitioner may miss important opportunities to think about what he is doing. He may find that...he is drawn into patterns of error which he cannot correct....A practitioner's reflection can serve as a corrective....Through reflection, he can surface and criticize the tacit understandings that have grown up around the repetitive experiences of a specialized practice, and can make new sense of the situations of uncertainty or uniqueness which he may allow himself to experience (p.61).

The interviews conducted for this study confirm the notion that reflection-in-action is an important part of the process through which instructional adaptation and change actually occur and that the reflective process is enhanced when it is done with the input of others. The instructional evaluation activities which were described as particularly useful for improvement and development purposes were consistently described as involving a discussant-critique component.

Integrated Critique

Within universities, the process of collective, cooperative reflection-in-action holds promise as a vehicle for instructional improvement far beyond the obvious benefits to individual faculty members and single courses. At any given hour in any given day, a single faculty member may be teaching a single course to a single group of students. All the same, these single courses are, or should be, related to one another through subject matter and/or programmatic and university instructional goals and objectives. Furthermore, they occur within a context of institutional values and physical and temporal arrangements. Quality university instruction is dependent, therefore, upon the ongoing questioning, discussion, and
examination of the effects these realities have upon individual courses and vice versa.

The institutionalized patterns of thought and action which manifest educating processes in universities do more than simply shape and constrain the behavior of individuals. They also manifest, in themselves, the "theories of educating" that are implemented in that particular setting. Such institutional arrangements, in most cases, are not based on anything other than institutional history. They, and many educational practices, are grounded upon traditions and conventions that have not been questioned or tested systematically. Yet they probably affect educational outcomes just as much, and maybe more than, any laws of nature underlying school learning (Sanders, 1982, p. 12)

Improved information, more thorough and complete understanding, and greater proficiency in systematically ascertaining whether proposed actions...will lead to attainment of desired goals is more likely to contribute to the improvement of the quality of teaching and the performance of educational systems than any other approach to the improvement of education...(Sanders, 1982, p. 19)

Organizational Rewards, Incentives, Norms

If faculty are interested in continuously improving their practices, and if interactive, collective reflection is an effective means of doing so, why does such a process not occur more often? The most obvious answer seems to be that there are few organizational rewards or incentives for doing so. A more complex, but equally significant reason is that many of our cultural values and norms discourage collective, critical analysis of professional and institutional practices.

Rewards, incentives. Although administrators in multiversities commonly claim that instruction is an activity highly valued by the organization, this is not the perception of many university faculty. The
participants in this study, like their counterparts at other universities (See Thorne et. al., 1976 and Boyd & Schientinger, 1976), repeatedly voiced the viewpoint that it is research, not teaching, that is rewarded. The rewards which the faculty refer to are promotion and tenure. In general, the faculty members in multiversities do not perceive it as being in their best interest to invest much effort in assessing and improving their instruction. While they acknowledge the importance of providing quality instruction, faculty feel that they can't afford to expend too much effort on their instructional activities—the payoffs are too small.

I had a computer-based instruction course I wanted to develop when I came here—but that has to go on the back burner because it takes a fair amount of development time, and unless the research associated with it is directly related to my discipline it won't get me very much in terms of recognition and rewards. It's a relatively low payoff. That may change when I get tenure. Those kinds of things may move more to the front a little bit more. (f03)

I'm not negative toward teaching in terms of importance. I have to be negative about it in terms of goal satisfaction or anything else. You are not going to get anything by teaching except perhaps to teach more which is going to put you that much further in the hole. Anything I've gotten here in the way of extra help I've bought and that comes from outside funding. That is research. There is no financial reward for teaching that I can see. (f32)

This perception appears to be reinforced at the department level through departmental performance evaluation practices. The faculty in this study perceived performance evaluation activities as being directed exclusively toward salary, promotion and tenure decisions. They also perceived research accomplishments as the focus during annual performance reviews. For the most part, they did not perceive department chairs or colleagues as
interested or involved in assessing and critiquing instructional performance
for the purpose of simply improving and developing instructional practices.
This perception appears to significantly, and understandably, influence a
faculty member's willingness and desire to ask for outside review.

An example of the effect that the multiversity's reward system has on
faculty attitudes and behavior is illustrated in the comments of an
associate professor:

My attitude has changed since I came here seven years ago. Prior to coming here I always wanted to be at the university. I found out the job was 75% teaching and 25% research. I always wanted to teach, although I had never done it. I thought: This is great. I'm going to get to do both research and teaching. Well, since I've been here, I'd just prefer to get rid of the teaching. It interferes with the things that bring the reinforcing. People don't ask me about my teaching— they ask about my research. When I have those two quarters I have to teach, I am locked in. I put too much time into it and it hurts my research program tremendously. I'm not able to stay up on proposals and the development of proposals and writing manuscripts. If I could reduce my teaching responsibility, I'd do it right away.

Don't misunderstand. I love teaching. If the university could convince me that they were sincere when they said, "we're going to make a serious evaluation of you as a teacher, and you're going to be rewarded for teaching"—they haven't done that up to now—I think it would be one of the greatest boons to this university that we have. (f25)

It has long been pointed out that, until teaching is rewarded within the multiversity, we cannot expect faculty to seriously examine and improve it. A recent report from the Study Group on the Conditions of Excellence in Higher Education (1984) has again restated this position. They offered,
among others, the following conditions as necessary for improving the support of teaching in U.S. colleges and universities:

* An institutional reward system that emphasizes effective teaching in faculty personnel decisions such as hiring, retention, tenure, promotion, and compensation;

* A multidimensional approach for evaluating college teaching and for helping faculty improve teaching that involves students, colleagues, administrators, and consultants who use a variety of methods and criteria in assessing instructional quality; and

* An environment that reinforces the commitment of faculty to teaching through contributions to the literature on college instruction and student development, proceedings of the teaching divisions of professional societies and higher education associations and the development of significant courses, curricula, and instructional materials such as textbooks and software.

The implementation of these recommendations would clearly support faculty involvement in collective critical analysis of individual and institutional instructional practices. It is proposed, however, that persons who might attempt to implement such recommendations would quickly encounter a set of held beliefs, cultural values, and norms that, if left unexamined, could easily sabotage such changes.

Beliefs, values, norms. The faculty who participated in this study expressed a variety of beliefs, or made statements that seemed to describe, norms which are operating in the multiversity and which serve as barriers to collective instructional critique. Some of these statements included the following:

-- You are very careful about poking your nose into troubles of peers; that's not your job.

-- We regard each other's ground in the classroom as sacred
We don't care to talk about the personal aspects of teaching.

--Faculty sitting in on one another's classes would not be taken too positively because it would be looked at like a policeman or an observer.

These beliefs seem to be derived from both the American culture at large and the university culture specifically. Independence, individualism, and personal freedom are philosophical tenets that are bred and nurtured in and through our cultural experience and institutions. One important enactment of these tenets in university culture is the concept of academic freedom. Academic freedom has at times been offered as a rationale by faculty for disallowing their instruction to be critiqued by others. The AAUP has defined academic freedom in the following way:

(a) The teacher is entitled to full freedom in research and in the publication of the results...

(b) The teacher is entitled to freedom in the classroom in discussing his subject, but he [she] should be careful not to introduce into [the teaching situation] controversial matter which has no relation to the subject.

(c) The college or university teacher is a citizen, a member of a learned profession, and an officer of an educational institution. When [the teacher] speaks or writes as a citizen, he [she] should be free from institutional censorship or discipline, but his [her] special position in the community imposes special obligations. As a [person] of learning and an educational officer, [the teacher] should remember that the public may judge his [her] profession and the institution by his [her] utterances. Hence [the teacher] should at all times be accurate, should exercise appropriate restraint, should show respect for the opinions of others, and should make every effort to indicate that he [she] is not an institutional spokesperson (AAUP, 1969, pp. 35-67).
Given this definition, cooperative, collective critical analysis of instructional practices aimed at improvement could be conducted without violating the intentions of this important principle. However, because evaluation techniques have sometimes been used by administrators as adversarial accountability tools, it is understandable why faculty might fear, and protect themselves and their students against, such tactics.

The non-interventionist stance expressed in the faculty comments on the previous page probably reflects the position assumed by the particular faculty member about his or her own instruction more than it reflects the viewpoint of all faculty. In fact, the majority of faculty claimed that they would like more colleague involvement in critiquing instructional practices.

--I'd like to see more people in the classroom.

--I'd like more evaluation by my peers.

--We should be more active in helping the people who are starting out...they should have a discussion of what's happening in the courses...It should be an ongoing thing.
--We could do each other a great deal of good by sitting in on each other's classes and sharing information.

--I would like to see peer review of courses if we are going to give more than lip service to teaching quality.

--We ought to have to evaluate other people who are teaching in the department.

--There is no defensible reason why you would not do peer review [of teaching]...If we can subject our written [research] work, I can't understand why we can't subject our teaching to a like kind of system....I don't know why we can't have people who are willing to look at a very detailed parallel between what we do for research and what we do for teaching. I know they aren't the same, but there are some generalizations that I'm sure will carry over.

--I think teaching requires every bit as much effort [as research]...The reason it [teaching] isn't held in such high esteem is that there is not review of it. Nobody knows, nobody critically evaluates you...The quality of teaching could probably be upgraded if we were really serious about peer reviews.

These statements may or may not describe how faculty would feel if colleagues or others actually became involved in a process of instructional critique. Regardless, this set of statements at least implies an openness to the process of critical review. Another, and perhaps more critical, belief or norm is more subtly expressed through these comments. Several of the statements reflect the perception that evaluation is something that is done to someone rather than with each other. It's an understandable impression given the history and current enactment of evaluation activities on university campuses.

Most formal evaluation activities in our universities have been externally driven. Student evaluation of teaching programs, accreditation
reviews, and faculty performance reviews have all been founded and operated on the notions of consumer and administrative accountability. They have been used as means of external monitoring and control. It therefore appears difficult for faculty to think within the constructs of evaluation without also thinking in terms of persons external in an event passing judgment on those within it. They have experienced few, if any, models for conducting constructive, cooperative, and collective critical reviews of personal and institutional practices.

A related institutional norm. Judging from these interviews, a related institutional norm is: Do not question or examine organizational arrangements beyond your own courses. This norm may be derived from the particular philosophies and politics of individual faculty, but it appears to be so universally and habitually followed that it seems more likely to be a reflection of the organizational structure. For example, there wasn't a faculty member who claimed to enthusiastically support the process through which promotion and tenure decisions currently are made, but there also wasn't a faculty member who was involved in trying to change current P&T practices. There were faculty who recognized the effects of institutional arrangements on the classroom, but they weren't attempting to communicate their concerns and to change things.

When we teach in the branches we teach two days a week. I go with no break. I finish up in two hours. I can cover more material, easier in those two days than I can in my classes that meet every day. But I don't think it is easier on the student. Two hours is a lot of time to ask a student to absorb a bunch of garbage. It's a lot easier to teach, but it's not best for learning. Q: Do you talk about those things? How instruction is organized, what is feasible to change?
A: No. Because I think we feel that is pretty much out of our control. I've given up. It's not important. The mountain has been here so long. I don't know where to start. (f44)

The size of the university makes it difficult for faculty to communicate and formulate their common concerns. They seem to experience a sense of alienation from one another and the university at large, and it appears that both formal and informal organizational arrangements contribute to this experience.

---We have never had a faculty lounge in this building; we are not encouraged to get together. (f14)

---I'd like to believe that our faculties need to get together on an informal basis more often because we suffer from the problem of size. I think I hark back to the time when the university had 10,000 students and faculties knew each other on a personal basis whether it was over lunch or socially. It doesn't occur on this campus. There are so many things that are done informally over a cup of coffee that don't happen here. I don't even belong to the Faculty Club. There was a time when that would have been ridiculous because it was the social center of the university, but it is so big now that I can go over there and not see four people I know. I think we suffer from size and the increase in size of the administration. We are so big, so diverse, that we don't hear each other very well. (f16)

---We only have faculty meetings about once a quarter and when we do we don't really discuss the curriculum. We don't ever have a seminar or retreat on teaching. There's very little interaction among faculty. My interactions during the day are with students and with people on the telephone. Literally people can be in this building and in this hallway for three or four days and I'll never have an occasion to interact with them. I interact with my graduate students, people out in the state, other students, etc. If faculty do talk, it's pretty much about the basketball game. You just don't interact that much. (f41)
Given the extent to which some faculty feel isolated and fragmented from the university at large, it may seem overwhelming for them to think in terms of restructuring the way instruction is viewed and evaluated within the multiversity. It might, however, be perfectly conceivable for them to think in terms of changing practices within their departments. If faculty were encouraged and supported to examine within their departments why and how they design, organize, implement and evaluate instruction within and across courses, institutional instructional practices could not long go unexamined. Over time, these institutions might themselves become collectively reflective.

A practitioner who reflects-in-action tends to question the definition of his task, the theories-in-action that he brings to it, and the measures of performance by which he is controlled. And as he questions these things, he also questions elements of the organizational knowledge structure in which his functions are embedded. In contrast, to the normal bureaucratic emphasis on uniform procedures, objective measures of performance, and center/periphery systems of control, a reflective institution must place a high priority on flexible procedures, differentiated responses, qualitative appreciation of complex processes, and decentralized responsibility for judgment and action. In contrast to the normal bureaucratic emphasis on technical rationality, a reflective institution must make a place for attention to conflicting values and purposes. But these extraordinary conditions are also necessary for significant organizational learning (Schon, 1983, pp. 337-338).
Summary.

Instructional decision-making in higher education is not an activity that lends itself to a purely technical-rational model of problem solving.

Technical rationality depends on agreement about ends. When ends are fixed and clear, then the decision to act can present itself as an instrumental problem. But when ends are confused and conflicting, there is as yet no "problem" to solve. A conflict of ends cannot be resolved by the use of techniques from applied research. It is rather through the non-technical process of framing the problematic situation that we organize and clarify both the ends to be achieved and the possible means of achieving them (Schon, 1983, p.41).

The field of educational and instructional evaluation has perhaps been too well grounded in the traditions of the rational model. We have often attempted to treat the acts associated with teaching, learning, and educating as if they were value-free, context-free activities that could be objectified, instrumentized, and quantified. We have sometimes behaved as though the "objective" data derived through our instruments would somehow produce all we would need to know in order to make "correct" choices about teachers, learners, and instructional activities. In higher education our use of student ratings of teaching is a poignant example of the inadequacies of a purely technical-rational model for educational decision-making.

Instruction is a complex human activity that occurs in a complex organizational context. It is laden with enormous social, psychological and moral significance. As such, it requires continual and collective inquiry, examination, and critique. If the experiences of the faculty who participated in this study are at all representative of the faculty at large, we should be seriously concerned that such purposeful inquiry is occurring
so infrequently and sporadically. We should also be cautious about blaming the faculty alone for current conditions.

There are several organizational conditions that contribute to, if not determine, the faculty member's approach to instructional inquiry. The university's current approach to performance and instructional evaluation discourages rather than facilitates faculty critique of their professional practices. The adoption of external accountability models and the centralization of personnel decision-making forces faculty to concentrate on collecting instructional data that will "count" and that can be counted during P&T decisions. University faculty are, overall, quick learners, and the most concrete reward system within the multiversity (salary, promotion, tenure) encourages faculty to minimize their involvement in instructional activities and to maximize their involvement in research activities. Finally, there are few formal or informal opportunities for faculty to discuss instructional issues and concerns. Even the physical arrangements of office space and classrooms encourage isolation and individualism rather than collaboration, cooperation, and critique.

Knowing-in-action, as described by Schon, is perhaps an unavoidable condition of functioning as a university faculty member, and individual reflection-in-action is a desirable and, perhaps, necessary condition for continuously adapting and improving the faculty member's instructional practices. However, reflection-in-action, which includes discussion and cooperative critique, allows us to raise the shades in the darkened rooms of our own, and the organization's beliefs and behaviors. It allows us to potentially form new, more refined, and more complete images of what we're doing, why we're doing it, and how we want to do it in the future. As a
process, collective reflection-in-action is not the answer to all the problems associated with adapting and changing our university's instruction—there is no single, now-and-forever-always, answer. Just the same, an as executive for Hewlett Packard, one of the exemplar corporations identified in *In Search For Excellence*, points out:

We're really not sure exactly how the innovative process works. But there's one thing we do know: the easy communications, the absence of barriers to talking to one another are essential. Whatever we do, whatever structure we adopt, whatever systems we try, that's the cornerstone...(Peters & Waterman, 1982, p. 218).
BIBLIOGRAPHY


165


Ecker, G. (1980). *Administration in higher education: Making the most of ambiguity*. Unpublished manuscript, The Ohio State University, College of Education, Columbus, Ohio.


The Ohio State University Faculty and Staff Directory. (1985). The Ohio State University, Columbus, Ohio.


Rules of the University Faculty. (1984). The Ohio State University, Columbus, Ohio.


### Percentage of Time Reported in Each Activity Area by Each Faculty Participant

#### Instruction Primary Activity

<table>
<thead>
<tr>
<th>Rank</th>
<th>I</th>
<th>R</th>
<th>PS</th>
<th>Adm.</th>
<th>Adv.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full n=8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>15</td>
<td>05</td>
<td>15</td>
<td>05</td>
<td>05</td>
</tr>
<tr>
<td>50</td>
<td>20</td>
<td>25</td>
<td>00</td>
<td>05</td>
<td>05</td>
</tr>
<tr>
<td>45</td>
<td>20</td>
<td>10</td>
<td>10</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>10</td>
<td>05</td>
<td>15</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>49</td>
<td>25</td>
<td>05</td>
<td>11</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>62</td>
<td>25</td>
<td>05</td>
<td>06</td>
<td>02</td>
<td></td>
</tr>
<tr>
<td>65</td>
<td>15</td>
<td>05</td>
<td>05</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>55</td>
<td>10</td>
<td>05</td>
<td>25</td>
<td>05</td>
<td></td>
</tr>
<tr>
<td>mean</td>
<td>63</td>
<td>18</td>
<td>8</td>
<td>11</td>
<td>8</td>
</tr>
<tr>
<td>Associate n=2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>55</td>
<td>05</td>
<td>10</td>
<td>15</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>70</td>
<td>05</td>
<td>05</td>
<td>05</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>mean</td>
<td>63</td>
<td>5</td>
<td>8</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Assistant n=5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>70</td>
<td>05</td>
<td>05</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>65</td>
<td>10</td>
<td>10</td>
<td>05</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>60</td>
<td>20</td>
<td>10</td>
<td>00</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>55</td>
<td>10</td>
<td>10</td>
<td>05</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>65</td>
<td>20</td>
<td>05</td>
<td>05</td>
<td>05</td>
<td></td>
</tr>
<tr>
<td>mean</td>
<td>63</td>
<td>13</td>
<td>8</td>
<td>5</td>
<td>11</td>
</tr>
</tbody>
</table>

Total n=15
### Research Primary Activity

<table>
<thead>
<tr>
<th>Rank</th>
<th>I</th>
<th>R</th>
<th>PS</th>
<th>Adm.</th>
<th>Adv.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Full</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n=4</td>
<td>30</td>
<td>50</td>
<td>10</td>
<td>00</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>40</td>
<td>20</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>60</td>
<td>15</td>
<td>10</td>
<td>05</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>50</td>
<td>00</td>
<td>10</td>
<td>15</td>
</tr>
<tr>
<td>mean</td>
<td>21</td>
<td>50</td>
<td>11</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td><strong>Associate</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n=1</td>
<td>25</td>
<td>50</td>
<td>10</td>
<td>05</td>
<td>10</td>
</tr>
<tr>
<td>mean</td>
<td>25</td>
<td>50</td>
<td>10</td>
<td>05</td>
<td>10</td>
</tr>
<tr>
<td><strong>Assistant</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n=4</td>
<td>35</td>
<td>50</td>
<td>05</td>
<td>05</td>
<td>05</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>55</td>
<td>05</td>
<td>00</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>55</td>
<td>04</td>
<td>08</td>
<td>08</td>
</tr>
<tr>
<td>mean</td>
<td>26</td>
<td>53</td>
<td>05</td>
<td>05</td>
<td>12</td>
</tr>
</tbody>
</table>

### Balanced Between Research and Instruction

<table>
<thead>
<tr>
<th>Rank</th>
<th>I</th>
<th>R</th>
<th>PS</th>
<th>Adm.</th>
<th>Adv.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Full</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n=3</td>
<td>35</td>
<td>40</td>
<td>10</td>
<td>05</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>30</td>
<td>00</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>46</td>
<td>00</td>
<td>02</td>
<td>02</td>
</tr>
<tr>
<td>mean</td>
<td>38</td>
<td>39</td>
<td>03</td>
<td>09</td>
<td>11</td>
</tr>
<tr>
<td><strong>Associate</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n=8</td>
<td>35</td>
<td>35</td>
<td>05</td>
<td>10</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>35</td>
<td>30</td>
<td>15</td>
<td>05</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>40</td>
<td>02</td>
<td>10</td>
<td>08</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>35</td>
<td>10</td>
<td>10</td>
<td>05</td>
</tr>
<tr>
<td></td>
<td>45</td>
<td>45</td>
<td>00</td>
<td>05</td>
<td>05</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>44</td>
<td>05</td>
<td>00</td>
<td>01</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>45</td>
<td>05</td>
<td>05</td>
<td>05</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>40</td>
<td>05</td>
<td>01</td>
<td>04</td>
</tr>
<tr>
<td>mean</td>
<td>42</td>
<td>39</td>
<td>05</td>
<td>06</td>
<td>07</td>
</tr>
</tbody>
</table>
(Continued)

Public Service Primary Activity

<table>
<thead>
<tr>
<th></th>
<th>Full</th>
<th>Associate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n=2</td>
<td>n=2</td>
</tr>
<tr>
<td></td>
<td>I</td>
<td>R</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>05</td>
<td>20</td>
</tr>
<tr>
<td>mean</td>
<td>13</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>05</td>
</tr>
</tbody>
</table>

Balanced Three Ways

<table>
<thead>
<tr>
<th></th>
<th>Full</th>
<th>Associate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n=1</td>
<td>n=1</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>35</td>
</tr>
</tbody>
</table>
Appendix B

Faculty Background Sheet

Faculty Background Information

1. Courses which you taught this academic year:
   Course Number (s)       Approximate Enrollment (s)

2. I am tenured: yes ______ no ______

3. Number of years as a college teacher: ______

4. Number of years at OSU: ______

5. The academic discipline from which you received your undergraduate degree:

6. The academic discipline (s) from which you received your graduate degree (s), including Masters and Doctorate:

7. Please indicate, by percentages, the approximate amount of your time spent in each of five university activity areas during this past academic year:
   Instruction ______%
   Research ______%
   Public Service ______%
   Administration ______%
   Student Advising ______%
   Total 100%
Appendix C

Definitions of Codes Used for Organizing Faculty Interview Data

.01 Student Evaluations: types of
- type(s) of student evaluations the faculty member reportedly currently uses.

.02 Informal feedback
- comments which describe casual, unofficial ways information is collected and/or used to make evaluative judgements about instructional practices.

The opposite of informal feedback--formal feedback--would be processes which use validated instruments and/or other means through which verifiable, empirical evidence about instructional practices is purposefully collected and used for the purpose of making evaluative judgements about instructional practices.

.03 Peer review
- comments on means through which other faculty do, or could, critique, observe, or otherwise provide input to the faculty member or others about the participating faculty member's instructional practices.

.04 Chair review
- comments on process(es) through which the faculty member's department chair does, or could, critique, observe or otherwise comment on the faculty's instructional practices.

.06 Reasons for using the Cornell
- If a faculty member reportedly uses the Cornell instrument, his or her self-reported rationale for using same.

.07 Negative SET comments
- comments made about the SET that could be interpreted as negative, i.e., disadvantages, problems, etc.

.08 Changes
- comments describing actual changes in instructional practices that faculty have made and their perceived reasons for making such changes. ("Instructional practices' includes such things as teaching style, content covered, tests, course sequence, etc.)
Analysis of Student Evaluations: procedure

- comments which describe the process through which the faculty member analyzes formal student evaluations of teaching.

Exit Interviews

- comments which refer to exit interviews as a source of evaluative information.

Performance Standards

- comments related to the perceived standards, criteria and/or clarity of expectations for acceptable performance within the university and/or individual departments.

Attitude toward teaching

- comments which might be interpreted as reflecting or revealing a particular faculty member's attitude toward or philosophy of teaching.

Weight given teaching

- comments related to the perceived relative weight given to teaching performance during promotion, tenure and/or salary decisions. Perceived value placed on teaching role by university, college, department and/or others.

Evidence for P&T

- comments related to evidence about performance that the faculty member either included in his or her P&T packet or reportedly witnessed as included in the packets of others.

Faculty development

- comments which describe activities in which the faculty member or others have, or could, be involved for the purpose of acquiring or improving teaching skills.

Rewards

- comments which referred specifically to "rewards" associated with teaching and which reveal faculty's held beliefs about internal versus external rewards for teaching.

Administrators' influence on faculty attitudes toward teaching

- comments related to the relationship between formal and informal administrative policy and faculty attitudes and behavior regarding teaching.
Assignments
- comments related to faculty "work" assignments.

Pattern of use
- comments which describe the faculty member's and or departments' use of student evaluations over time.

Reasons for using SETs
- if a faculty member reportedly uses SETs, his or her self-reported rationale for using same.

Misc. Instructional Evaluation
- comments describing means used to evaluate instructional practices other than those processes described under .x01, .x02,

Curriculum Development
- comments which describe the process(es) through which courses are designed and developed, i.e., comments about who decides, and/or approves, and/or reviews subject matter, materials, testing procedures and policy, and the process(es) through and frequency with which those decisions are addressed.

Reasons for using student evaluations
- comments which provide espoused rationales for conducting written student evaluations—rationales not listed under .x06 or

Alumni
- comments that refer to using alumni as a source of evaluative information about instructional practices.

Recommendations
- comments which recommend changes in current instructional evaluation practices.

Attitude toward student evaluations
- comments which reflect faculty attitudes toward using students as a source of evaluative information about instructional practices.

P&T Process: feedback
- comments which refer to how faculty are provided performance information, and/or status reports about P&T actions, through the P&T process.
.x36 P&T Process: documentation

- comments referring to the packets of information about faculty performance which are compiled for use during the P&T process.

.x37 Aspects of Instruction to focus on

- faculty's descriptions of types of questions they ask students to respond to during course evaluations; the instructional factors about which they are reportedly interested in soliciting feedback.

.x38 Negative Cornell

- comments about the Cornell Diagnostic Instrument that could be interpreted as negative, i.e., disadvantages, problems, etc.

.x39 Accreditation Reviews

- comments referring to reviews conducted by external accrediting bodies.

.x40 Changes in P&T Process

- perceived changes in the university's and/or department's promotion and tenure process.

.x41 Policy

- comments referring or related to university and/or department policy regarding conducting student evaluations of teaching.

.x42 Student evaluations: actual use

- comments which describe or relate to how, for what purpose, and by whom the information collected through written student evaluations is actually used.

.x43 Annual Reviews

- comments describing the annual review process as, or if, practiced in departments and its relationship to evaluating instructional practices.

.x45 Written Comments

- comments describing the use and perceived value of students' written comments about instructional practice(s).
Weight given public service

- comments related to the perceived relative weight given to public service role during promotion, tenure and/or salary decisions. Perceived value placed on public service role by university, college, department and/or others.

University Program Review

- comments referring to the university's formal academic program review process.
Appendix D

Sample Page of Computer Printout for Organized Test Categories

<table>
<thead>
<tr>
<th>Line Nos.</th>
<th>Text</th>
<th>Coding</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.</td>
<td>STUDENT EVALUATIONS: TYPES OF</td>
<td>.01 F01 0731 119</td>
</tr>
<tr>
<td>4.</td>
<td>I ALWAYS CONDUCT STUDENT EVALUATIONS OF MY COURSE. I HAVE MY OWN FORM.</td>
<td>.01 F01 0731 119</td>
</tr>
<tr>
<td>5.</td>
<td>STANDARD QUESTIONS THAT I ASK: WHAT DO YOU LIKE LEAST ABOUT THE COURSE?</td>
<td>.01 F01 0731 119</td>
</tr>
<tr>
<td>6.</td>
<td>THERE THAT I COVERED THAT I SHOULD NOT COVER</td>
<td>.01 F01 0731 119</td>
</tr>
<tr>
<td>7.</td>
<td>QUESTIONS OVER THE LAST 7 OR 8 YEARS. FOR TWO OR THREE YEARS THERE WAS A FAIRLY STANDARD UNIVERSITY FORM THAT I USED—AT LEAST MACHINE GRADED. I USED THAT FOR ABOUT 3 YEARS. THE INFORMATION SEEMED TO HAVE DISAPPEARED.</td>
<td>.01 F01 0731 119</td>
</tr>
<tr>
<td>8.</td>
<td>AND I TOOK A LOOK AT THE CORNELL FORM, AND I DIDN'T THINK THAT WAS APPROPRIATE FOR ANYBODY'S COURSES.</td>
<td>.01 F01 0731 119</td>
</tr>
<tr>
<td>9.</td>
<td>I'VE GENERALLY USED THE CORNELL ON OCCASION IF I DON'T MEET DEADLINES SOMETIMES I'VE MADE UP ONE OF MY OWN.</td>
<td>.01 F02 0631 050</td>
</tr>
<tr>
<td>10.</td>
<td>STUDENT EVALUATIONS: TYPES OF</td>
<td>.01 F02 0631 050</td>
</tr>
<tr>
<td>11.</td>
<td>USE THE CORNELL AND ONE OF MY OWN FOR EACH COURSE. USE A VERY DETAILED, SELF-DESIGNED QUESTIONNAIRE FOR EACH COURSE I'M DEVELOPING.</td>
<td>.01 F02 0631 050</td>
</tr>
<tr>
<td>12.</td>
<td>THE FIRST TIME I TAUGHT I USED A RATHER ELABORATE FORM. I ACTUALLY HAD THE INFORMATION KEY PUNCHE D AND DID A STATISTICAL ANALYSIS. I GOT THE QUESTIONS FROM MY OWN FORM PARTLY FROM AN INSTRUCTOR WHO HAD BEEN TEACHING THE COURSE FOR ABOUT TEN YEARS AND I WANTED TO CONTINUE TO ASK THE SAME QUESTIONS AND SOME FROM OTHER SOURCES.</td>
<td>.01 F03 0112 126</td>
</tr>
<tr>
<td>13.</td>
<td>WE USE A SET FORM THAT WAS PUT TOGETHER BY THE DEPARTMENT. WE ALL USE THE SAME ITEMS. A GRADUATE STUDENT COMES IN TO ADMINISTER IT.</td>
<td>.01 F04 0421 018</td>
</tr>
<tr>
<td>14.</td>
<td>AS FAR AS STUDENT COMMENTS THAT THEY PROVIDE, DID YOU SAY THAT FREQUENTLY STUDENTS COMMENT ON THE BACK OF THIS FORM—THE DEPARTMENT SET FORM?</td>
<td>.01 F04 0421 018</td>
</tr>
<tr>
<td>15.</td>
<td>TAS WILL ADMINISTER THE FORM. THE FORM WILL SAY THEY HAVE A RIGHT TO COMMENT—ENCOURAGE IT. Q: DO YOU KNOW ABOUT WHAT PERCENT RIGHT</td>
<td>.01 F04 0421 018</td>
</tr>
</tbody>
</table>