

DEPARTMENTAL EXCELLENCE: CRITERIA USED BY
DEPARTMENT HEADS AT PUBLIC COMPREHENSIVE
UNIVERSITIES AND COLLEGES

By

APIWAT WATANANGURA
"

Bachelor of Education
Chulalongkorn University
Bangkok, Thailand
1979

Master of Education
Loyola College
Baltimore, Maryland
1981

Submitted to the Faculty of the Graduate College
of the Oklahoma State University
in partial fulfillment of the requirements
for the Degree of
DOCTOR OF EDUCATION
December, 1984

Thesis
1984D
W324d
cop. 2



DEPARTMENTAL EXCELLENCE: CRITERIA USED BY
DEPARTMENT HEADS AT PUBLIC COMPREHENSIVE
UNIVERSITIES AND COLLEGES

Thesis Approved:

John J. Hardin
Thesis Adviser

Robert B. Gamm

Ronald Beer

William E. Leggett

Norman D. Murken
Dean of the Graduate College

ACKNOWLEDGMENTS

The research and writing of this dissertation came about with the help and assistance of others. The writer wishes to express his gratitude and appreciation to Dr. John J. Gardiner for his excellent direction and guidance throughout the doctoral experience. Appreciation is also extended to Dr. Robert B. Kamm, committee chairman, for his support, understanding, and guidance, and to Dr. Lynn K. Arney and Dr. William E. Segall for their valuable contributions as members of the dissertation committee. Very special thanks to Dr. Ronald F. Beer, who replaced Dr. Lynn K. Arney in the final stages of the dissertation process.

A special word of gratitude is offered to Dr. Jo Campbell for her assistance in data analysis and to Dr. and Mrs. M. Jesada for their constant moral support and encouragement. Special recognition and deep gratitude are extended to the writer's parents for their endless and tireless support across the sea throughout the writer's four and a half long years in America.

TABLE OF CONTENTS

Chapter	Page
I. INTRODUCTION.	1
Need for the Study	2
Problem Statement.	4
Definition of Terms.	5
Assumptions.	6
Limitations.	7
II. REVIEW OF RELATED LITERATURE.	8
The Academic Department.	8
Academic Department Heads.	14
Criteria for Determining Departmental Excellence:	
Graduate Level	18
Criteria for Determining Departmental Excellence:	
Undergraduate Level.	27
Comprehensive Colleges and Universities.	30
Origin.	30
Concept of Comprehensive Universities and	
Colleges.	32
Recent Studies.	36
Conclusion.	37
Summary.	38
III. METHODOLOGY	41
Design of the Study.	41
Description of the Population and Sample	41
Description of the Survey Instrument	42
Procedures, Including Data Collection.	43
Data Analysis.	44
IV. PRESENTATION AND ANALYSIS OF DATA	47
Pilot Study.	47
Response Rate.	49
Ranking of the Key Components.	51
Criteria of Departmental Excellence.	52
Similarities and Differences in Criteria Used for	
Judging Departmental Excellence Among Eight	
Disciplines.	56

Chapter	Page
Levels of Acceptability.	61
Summary.	63
V. FINDINGS, CONCLUSIONS, AND RECOMMENDATIONS.	65
Findings	67
Conclusions.	71
Recommendations for Further Study.	72
Concluding Thoughts.	73
Academic Departments.	73
Student Evaluation of Faculty Performance on Departmental Courses.	75
BIBLIOGRAPHY	78
APPENDIXES	84
APPENDIX A - BIGLAN'S TABLE	85
APPENDIX B - QUESTIONNAIRE COVER LETTER	87
APPENDIX C - QUESTIONNAIRE.	89
APPENDIX D - FOLLOW-UP LETTER TO NONRESPONDENTS	97
APPENDIX E - TABLES SHOWING RANKING OF KEY COMPONENTS	99

LIST OF TABLES

Table	Page
I. Ranking of Key Components (Pilot Study)	49
II. Substitutions of Departments Within Disciplinary Areas. .	51
III. Number of Responses in Each Discipline.	52
IV. Ranking of Key Components	53
V. Listing in Rank-Order of the Percentages of Respondents Designating Each Characteristic as Very Important, Mod- erately Important, and Unimportant.	55
VI. Ranking of Key Components in Eight Disciplines.	57
VII. Listing of Rank-Order (1-5) of the Percentages of Respond- ents of Each Discipline Designating Only as Very Impor- tant Criteria.	59
VIII. The Optimal and the Minimum Level of Performance of Each of the Five Most Important Criteria	63
IX. Ranking of Key Components in Pure-Hard Nonlife System Discipline.	100
X. Ranking of Key Components in Pure-Hard Life System Disci- pline	100
XI. Ranking of Key Components in Pure-Soft Nonlife System Discipline.	101
XII. Ranking of Key Components in Pure-Soft Life System Disci- pline	101
XIII. Ranking of Key Components in Applied-Hard Nonlife System Discipline.	102
XIV. Ranking of Key Components in Applied-Hard Life System Discipline.	102
XV. Ranking of Key Components in Applied-Soft Nonlife System Discipline.	103

Table	Page
XVI. Ranking of Key Components in Applied-Soft Life System Discipline.	103

CHAPTER I

INTRODUCTION

The term "excellence," like the concepts of truth, beauty, and nobility, is difficult to define. Once defined, however, scholars in education seemed to agree that the central meaning of the concept was its state of being "superior to other, first class, top quality" (Alley, 1982, p. 101), a state of being in which "only the few rather than many can obtain" (Cartter, 1966, p. vii). In presenting a report on excellence in education in America, the National Commission on Excellence in Education (1983) defined excellence at three levels:

At the level of individual learner, excellence means performing on the boundary of individual ability in a way that tests and pushes back personal limits, whether in school, in the workplace, or in learning as an adult. Excellence characterizes a school or college that sets high expectations and goals for all learners, then tries in every way to help students reach them. Excellence characterizes a society that has adopted these policies, for it will be prepared through the education and skill of its people to respond to the challenges of a rapidly changing world (p. 8).

Excellence in education created an "imaginary dilemma" (Gardner, 1961, p. 90). Arguments over quality versus quantity education, or elite versus mass education, as Gardner saw them, were "heated and rather pointless" (p. 89), because they were argued from different points of view. The twin goals of American education were to pursue excellence in education while seeking to educate every individual to the limit of his or her ability. Alley (1982) supported this notion

of excellence when he noted that "most students will achieve excellence in a relative sense, while a few will achieve excellence in the absolute sense--and that's about as excellent as you can get" (p. 108).

Need for the Study

In determining excellence in higher education, some studies have used criteria that were not related to academic departments. Such criteria included institutional resources and student achievement (Brown, 1967; Krause and Krause, 1970). Realizing, however, that academic departments were the institutional arm of the discipline, most studies, especially at the graduate level, used the academic department as a unit of analysis in determining institutional excellence (Trow, 1977).

It was agreed among scholars that using the academic department as the unit of analysis was appropriate because the department was the heart of the institution. It was the organizational unit that made the institution run; all administrative decisions were made at the departmental level (Heimler, 1967). Corson (1975) noted:

The department--is the basic organizational building block of a college or university. The work for which the institution exists is carried out in a principal part through the departments. It exerts a major influence on decisions that determine the character of the institution. . . . (p. 250).

Waltzer (1975) believed that the success of the institution depended upon the success of its academic departments. As a result, one found a logical agreement that saw departmental excellence as the way to determine institutional excellence.

There has not yet been an agreement among scholars in higher education as to which criteria should be used to determine departmental excellence. Earlier reputational studies of graduate education have used such criteria as quality of the program to prepare doctoral candidates (Hughes, 1925, 1934); quality of the program and quality of the faculty (Keniston, 1959; Cartter, 1966; and Roose and Andersen, 1970). Counting the number of citations and faculty publication rates were widely used in determining departmental excellence (Smith and Fiedler, 1971). The validity of each of these criteria was debatable (Astin and Solmon, 1981).

While some scholars believed that there was a real set of standards "on which all brand of excellence can be compared, others insist the system is too diverse, too large, and too complex to allow one final analysis of excellence" (Washburn, 1980, p. 32). According to Washburn, faculty members at Stanford were more concerned about what the standards of departmental excellence were than they were about who should set the standards. They believed, however, that each interest group should have its own standards. This was a concept that Gardner (1961) called individual differences. As one individual was different from the other, one type of institution was different from another type. Therefore, criteria for determining departmental excellence that were used at research universities might not be appropriate for other types of institutions. A person needed to know the highest standards and the best methods in his or her own job, so that it might be performed with "professional pride and a sense of craftsmanship" (Harris, 1981, p. 15). Likewise, not until each type of institution

had identified its own set of criteria for determining departmental excellence could one justly assess the achievement.

Most studies in institutional excellence focusing on the academic departments conducted so far were concentrated on well-known research universities. Lawrence and Green (1980) emphasized that the assessment of departmental excellence at the master's degree programs and the undergraduate level had been ignored. This was because the master's degree programs were usually considered a requirement toward the doctoral program, while undergraduate education was diverse in terms of scope and function. Comprehensive colleges and universities, according to the Carnegie's classification, offered the combination of a liberal arts program and professional program, usually not higher than the master's level. Studies on this subject had been lacking. There had been no studies conducted regarding the identification of criteria used in determining departmental excellence in comprehensive colleges and universities.

Problem Statement

The purpose of this study was to answer the following research questions:

1. Given five key components: community and university service, faculty scholarship and research, graduate program, students, and undergraduate teaching, how did department heads at public comprehensive colleges and universities rank these components in relative importance?
2. What were the criteria that department heads at public comprehensive colleges and universities used in determining departmental excellence?

3. Were there any similarities and/or differences in the rankings of key components and in criteria used for determining departmental excellence across disciplines in randomly selected public comprehensive universities and colleges in the United States?

4. What were the levels of acceptability of the criteria that were identified as very important in determining departmental excellence?

Definition of Terms

For the purpose of this research study, the following definitions were used:

Criteria were "those few properties, things, or events singled out for purposes of evaluating excellence or merit" (Washburn, 1980, p. 10).

Excellence was the state of being better when certain criteria regarding two or more things were compared.

Departmental Excellence was the state that each academic department or discipline strove with the best of its caliber to achieve its stated departmental mission.

Academic Department was a division in a college which is usually responsible for instruction, research, and services within a specific discipline.

Department Head referred to the person designated by the university as the official administrator of the department. The terms "department head," "department chairman," and "department chairperson" were used interchangeably.

Comprehensive Universities and Colleges I were the institutions that

offer a liberal arts program and at least two professional courses of study, such as engineering and business. All of them have either no doctoral program or else an extremely limited one; most award master's degrees (Levine, 1978, p. xxiv).

For purposes of this study, required full-time equivalent student enrollments were at least 1,000.

Comprehensive Universities and Colleges II were those institutions offering a liberal arts program with at least one professional course of study. For purposes of this study, required enrollments were at least 1,000 (Carnegie Council on Policy Studies in Higher Education, 1976).

Public Comprehensive Universities and Colleges were those type 1 and 2 universities and colleges identified by the Carnegie's classification of institutions of higher education. They included 354 public-supported comprehensive universities and colleges of both types (I and II).

Levels of Acceptability were the benchmarks on a numerical scale determined by value judgments to specify levels of performance in terms of optimal, unacceptable, and acceptable (Washburn, 1980).

Assumptions

The study was based on the following assumptions:

1. Department heads in the same discipline used the same criteria in determining departmental excellence. As the selected representatives of each discipline in the comprehensive universities and colleges, the criteria they might select as measures of departmental

excellence could represent the judgment of their discipline as a whole.

2. Department heads participating in this study truthfully identified criteria they actually used or have actually used in determining departmental excellence.

Limitations

There were two major limitations in this study:

1. There are 354 public comprehensive universities and colleges in the United States. In order to meet financial resources available for the study, random sampling was restricted to the acceptable minimum sample size suggested by Gay (1981).

2. The generalizability of the findings was limited to public comprehensive universities and colleges in the United States.

CHAPTER II

REVIEW OF RELATED LITERATURE

The Academic Department

Early American colleges were not departmentalized. The trend toward specialization in college and university curriculum, needs of the students, and increase in enrollments were the forces behind the development of the academic department in the American organization of higher education (Dressel and Reichard, 1970).

Departmentalization became necessary in those early years when it was impossible for one tutor to teach a single class in all subjects. In the second quarter of the nineteenth century, the beginning of departmental organization was easily discernible at both Harvard and the University of Virginia, stated Brubacher and Rudy (1976). By 1767, there were four departments at Harvard: Latin, Greek, logic and metaphysics, and mathematics and natural philosophy. However, a more formal recognition of departments came nearly a century later, noted Dressel, Johnson, and Marcus (1970). In 1825, Harvard was reorganized into six departments, a result of a student rebellion in 1823.

During the same period of time the University of Virginia began its instruction, and was organized into eight schools headed by professors (Dressel, Johnson, and Marcus, 1970). These schools were essentially the equivalents of the departments. In 1826, a year later, James March, the president of the University of Vermont,

divided his university into four departments, and permitted students who were not seeking degrees to pursue the studies of a single department (Rudolph, 1962).

During the 1880s, Cornell and Johns Hopkins were very successful in establishing autonomous departments, but "the real solidification of departmental structure and the academic rank system came in the 1890s" (Dressel, Johnson, and Marcus, 1970, p. 4). The authors further reported:

Harvard moved decidedly toward departmentalization about 1891-1892. Columbia was thoroughly departmentalized by the late nineties, with Yale and Princeton only somewhat slower in adapting this organization style (p. 4).

The department, then, was as much an organizational as an intellectual necessity; an efficient unit for making decisions about the curriculum, student careers, and the appointments and the promotions of staff--decisions that could no longer be made effectively by the president (Trow, 1977). Millett (1962) noted that:

Under the guidance or leadership of a chairman or executive officer, each department has a number of vital decisions to make. Ordinarily it is the department as a group which decides the general scope and specialization of a subject matter to be undertaken in the course offerings. Ordinarily it is the department which determines the individual member who shall be invited to join the group, within the staffing limits established by the dean or the president of the college or university (p. 83).

Furthermore, Dressel and Reichard (1970) observed:

. . . it soon became apparent that the reputation of a university depends upon the reputation of its departments and the scholars within them. Autonomy in the development of a department became a necessity if the university was to achieve a national reputation (p. 387).

The academic department was not universally accepted as the best of all possible modes of academic organization (McHenry, 1977).

However, on many campuses, it played an important role in determining action on personnel, curriculum, and research facilities (Brubacher and Rudy, 1976). In addition, Millett (1962) supported that:

In every college or university the customary first grouping of faculty members is the department. It is the department which brings together all persons with a common subject-matter interest. It is the department which expresses the common professional allegiance of the faculty (pp. 82-83).

He further stated that though there were alternatives to the academic department, such as divisions of related discipline or schools and colleges without formalized subdivisions, the fact remained that the academic department constituted the prevailing pattern of organization for the planning and management of the learning process and of research, creative activities, and public service.

Waltzer (1975) stated that the academic department was "where the action is" (p. 4). It was the key unit for the academic, as was reflected in many missions. They comprised the following:

1. instruction and advising of undergraduate majors;
2. instruction of undergraduate nonmajors;
3. instruction of graduate students;
4. postdoctoral fellowships;
5. advising or consulting with professors from other disciplines;
6. basic research;
7. applied research;
8. promoting discipline within a university;
9. promoting departmental views and interests within a university;

10. promoting the discipline and professionalism nationally;
11. exploring interfaces of the disciplines;
12. attaining national recognition for the department;
13. consultation services to business and industry;
14. consultation services to governmental units;
15. provision of scholarly and congenial environment; and
16. provision of a social and recreational network for those affiliated with the department.

At any rate, though academic departments varied in their missions, they were primary management units of colleges or universities, stated Millett (1979). Each unit determined the work plan of a group of faculty and other resources to carry out desired work plans. The scope of departmental management included departmental planning, academic affairs, faculty affairs, student affairs, budgetary affairs, and the evaluation of departmental performance.

Tucker (1981, p. 42) stated that "there has been no better way to organize colleges and universities than through academic departments based on recognized disciplines," which was similar to Churchill's famous statement about parliamentary government: "The academic department may be the worst way to organize knowledge, to administer its disbursement, and to create new knowledge, but it is better than any other way we have thus far devised" (p. 63).

Biglan (1973a, p. 195) stated: "One of the most easily overlooked facts about university organization is that academic departments are organized according to subject matter." In his study, multidimensional scaling was performed on scholars' judgments about the similarities of the subject matter of academic areas. One hundred

and sixty-eight scholars at the University of Illinois made judgments about 36 areas; 54 scholars at a small western college judged similarities among 30 areas. The purpose was to identify the same dimensions used by scholars at both of these institutions in order to determine general and important characteristics of academic areas.

The common dimensions of academic subject matter were perceived by scholars in both of these institutions. The most prominent dimensions distinguished hard sciences, engineering, and agriculture from social sciences, education, and humanities. Biglan (1973a, p. 201) labeled this dimension "hard-soft." This taxonomy assumed the significance of Kuhn's (1962) paradigm. The second dimension was the concern of the area with application to practical problems. Education, engineering, and agricultural areas were distinguished from hard sciences, social sciences, and humanities. The last dimension was the concern with life systems. Scholars distinguished biological and social areas from those that dealt with inanimate objects.

In his subsequent study, Biglan (1973b) constructed a table showing a cluster of academic areas in three dimensions (Appendix A). Depending on the characteristics of their areas, the author further stated that scholars differed in: (1) social connectedness among faculty members, (2) their commitment to teaching, research, administration, and service, and (3) scholarly output.

The Biglan model led to several studies. Smart and Elton (1975) tested the validity of the model by examining the consistency of the department heads' responses toward 11 goals with Biglan's three dimensions. Department heads numbering 1,646 in 32 state-supported universities which awarded doctorates and had enrollments of between 9,000

and 21,000 in 1972 were assigned to eight groups in Biglan's three dimensional model based on their departmental affiliations. They were asked to indicate the amount of emphasis their departments placed on each of the 11 departmental goals using a five-point scale. The multiple discriminant analysis procedures were used. Three discriminant functions were found to be statistically significant. The nature of the differentiation provided by these three discriminant functions was found to be highly consistent with the three dimensions in the Biglan model. The results suggested that the Biglan model had "considerable potential for providing a conceptual framework to guide future research investigating academic departments" (Smart and Elton, 1975, p. 580).

Smart and Elton (1976) conducted another study using data about the duties performed by department chairpersons in 32 state-supported universities. Chairpersons were asked to indicate the number of hours per week they spent on 27 duties. The results indicated the differences in the clusters, and the Biglan model was validated.

Creswell, Seagren, and Henry (1979) asked 120 department chairpersons at one major university to check on a trichotomous scale (none=1, some=2, considerable=3) to what extent they perceived a need for professional development on 99 tasks. The analysis procedures included the descriptive analysis of the tasks where a large percentage of the chairpersons indicated "some" or "considerable." The test for the similarity or difference of responses between the state college and the university was conducted by using the one-way analysis of variance. The researchers, then, classified the departments into the Biglan model. The differences in perceived need for professional

development for department chairpersons classified as hard or soft, pure or applied, and nonlife and life were analyzed by using three separate multiple discriminant analyses.

Results showed that department chairpersons had different professional development needs when grouped into three dimensions. Creswell, Seagren, and Henry (1979) concluded that the validation

. . . not only presented additional evidence to support the theoretical model that distinct academic environments exist on college campuses, but also advanced a technique for clustering department chairs for in-service training (p. 236).

Academic Department Heads

The early studies of quality or excellence of institutions of higher education were done by Hughes and Keniston (as cited in Cartter, 1966). Using the department as a unit of analysis, both studies ranked universities in several graduate disciplines. Both studies, according to Cartter, received such criticisms as geographical bias and rater bias from some critics. The only difference between the two studies was the choice of raters. While Hughes asked faculty at Miami University in Ohio to name distinguished scholars in their disciplines to serve as raters (Koelsch, 1981), Keniston asked department heads at 25 institutions to identify the top five departments in their fields (Keniston, 1959). As far as the Keniston study was concerned, to some critics, department heads were not the best judges of quality. They were not necessarily the most distinguished individuals in the departments and were atypical of their faculty members in terms of age, rank, and knowledge. They also tended to be older and more conservative

than were the faculty members (Cartter, 1966). Lastly, they tended to be "traditionalists" who had outdated perceptions.

Many studies that were related to the assessment of quality in higher education used other constituencies in the academic community rather than the department heads. They used faculty members, students, and alumni as raters, participants, and informers. For example, Hartnett, Clark, and Baird (1978), in rating the reputations of the doctoral programs (chemistry, history, and psychology), sent the questionnaires to the department heads and requested that they be completed by faculty members with graduate level responsibilities. Baird (1980) used the correlational approach to study the relationship between ratings of graduate departments and faculty publication rates. The subjects of this study were: 511 chemistry, 584 history, and 598 psychology faculty members in a national sample of doctoral programs; 791 chemistry, 893 history, and 967 psychology graduate students; and 430 chemistry, 349 history, and 393 psychology doctoral recipients. Washburn (1980) asked faculty members at Stanford University to identify criteria that could be used in judging departmental excellence.

As to the question of why the choices of raters or participants of these studies were what they were, those researchers did not explain. However, some speculations could be made. First, the researchers wanted to obtain as much information as they could from everyone involved. Hensley (1980) observed three distinct rater groups who participated in various studies in the period of over 50 years. The first rater group was a group of faculty who participated in the studies of Hughes (1925) and Fretter (1977). Keniston (1959) and Margulies (1973) used a group of administrators, and Cartter

(1966), Roose and Andersen (1979), and Clark, Hartnett, and Baird (1976) used an eclectic group.

Second, it was possible that role ambiguity and role conflict might affect the choice of raters not to include department heads. Leslie (1973) stated that the role of department head was a very ambiguous one. Unlike other positions in the academic community, the role of department head was ill-defined (Falk, 1980). Lee (1972) asserted:

In his own eyes he is still primarily a teacher who has assumed certain administrative tasks and responsibility. He has not, as it were, 'sold out' completely to the other side by becoming a dean. He is, therefore, quite often in conflict as to whether his role is one of spokesman for his colleagues in the department or whether it is one of an administrator who must make the decisions not only for the welfare of his department but for the welfare of the college and university as a whole (p. 54).

Role conflict could be defined as the incompatibility of expectations that one individual had toward another (Carroll, 1976). In the case of the department head, there was not just one individual's expectations. There were several others, as Brann (1972) illustrated:

The department chairman is caught between students who want a relevant education and sense they are being short-changed, faculty who believe he should provide them with ever-increasing salaries, decreasing teaching loads, and such benefits as secretaries, space, books, and travel funds, and above him is a dean and a central administration who want every penny pinched and accounted for and who produce a myriad of rules and regulations which limit the chairman's flexibility and options (p. 6).

Third, it was also possible that those researchers were well aware of the Keniston (1959) study criticisms. However, Cartter's (1966) findings might help argue those criticisms. In 1966, Cartter surveyed 100 institutions using eclectic group raters. The raters were asked to rate the graduate departments in those institutions

based on two criteria: (1) quality of the faculty and (2) quality of the program (Cartter, 1966). This study included 900 department heads, 1,700 distinguished senior scholars, and 1,400 carefully selected junior scholars. The comparison of the responses of the three subgroups dispelled those criticisms. The department heads rated the departments as well as other scholars did. However, it was found that they tended to be traditionalists in their perceptions and to rate the high-ranking departments slightly higher, and the low-ranking departments slightly lower, than other scholars. Dent (1978) conducted the reputational ratings using 1,005 department heads in 15 liberal arts disciplines at 67 universities and found the same halo effect as Cartter did. As a matter of fact, those studies that used faculty members, students, or alumni also reflected the same kind of halo effect and rater bias. There was no reason to believe that the use of subjects other than department heads would produce more validity and reliability of judgments than the use of department heads alone. Cartter concluded that if he had used only the department heads as raters, the results would not have been greatly different.

In conclusion, despite the ambiguity and conflict of the role, "the department chairman is the person who makes the institution run" (Brann, 1972, p. 6). He or she is the leader of the department who sets the pace and maintains standards (Mevey and Hughes, 1952). Realizing that the role of department head is extremely difficult, one who chairs it must have great talents in teaching and administration (Dressel et al., 1970). Gross and Grambsch (1968) surveyed faculty and administrators at 68 public and private universities and found that the department heads as a group were perceived as having

considerable say in decision-making. They usually regarded themselves as faculty.

Criteria for Determining Departmental Excellence:
Graduate Level

Though assessments of departmental excellence have been attempted since 1924, Cartter's An Assessment of Quality in Graduate Education in 1966 and its replicating follow-up study conducted by Roose and Andersen in 1970, both sponsored by the American Council on Education (ACE) have been considered "a standard benchmark within the academic community" (Drew and Karpf, 1981, p. 305). Essentially, distinguished scholars were asked to rate academic departments in their own discipline using two criteria: (1) quality of the faculty and (2) quality of the graduate program--accessibility of faculty to students, nature of curriculum, and quality of students (Cartter, 1966; Roose and Andersen, 1970).

Using rating as a guide in recruiting faculty at Miami University in Ohio in 1924, Hughes (as cited in Koelsch, 1981) asked faculty at Miami University to name 40-60 scholars in their disciplines to serve as raters of graduate departments. His study ranked 38 universities in 20 graduate disciplines. By 1934, the number of institutions awarding the doctorate had increased to 106, and a second Hughes (as cited in Lawrence and Green, 1980) study rated 59 universities in 35 fields as "adequate" or "distinguished" using adequate facilities for work in the various fields to prepare candidates for doctoral degrees as criterion.

The first post-war study to update Hughes' work was conducted in 1957 by Keniston (1959). The purpose was to determine the present reputations of various departments which offered programs leading to the doctorate at the University of Pennsylvania. Raters, who were department chairpersons at 25 institutions selected from the institutional members of the American Association of Universities, were asked to identify the top five departments in their fields based on a combined measure of doctoral program quality and faculty quality. The findings were then compared to Hughes' results to determine changes over a generation.

In the introduction of the 1966 study, Cartter (1966) noted that both the Hughes and Keniston studies had two major weaknesses: (1) geographical bias and (2) rater bias. In the Hughes study, geographical bias was almost inevitable because the most distinguished universities at that time were centered in the Northeast and the Midwest. Some critics, according to Cartter, commented that Hughes and Keniston failed to anticipate that raters would tend to overrank their alma maters. The author further commented on another flaw--especially with the Keniston study--the choice of department chairpersons as raters. To some critics, chairpersons were not necessarily the most distinguished individuals in the departments; they were atypical of their colleagues in age, rank, and knowledge and tended to be older and more conservative than were their peers. As a result, their perceptions were outdated.

With the experiences of earlier studies, Cartter (1966, p. 12) then surveyed 100 institutions forming the Council of Graduate Schools in the United States in 1961, as well as six universities that had

granted "100 or more doctorates (spread over three or more fields) in the preceding decade." In short, 1,663 graduate programs in 29 disciplines were rated. Four thousand survey respondents included department chairmen, distinguished senior scholars, and knowledgeable junior scholars who had completed their formal training not more than 10 years earlier. They were asked to rate academic departments based on two criteria: (1) quality of the faculty and (2) quality of the graduate program. The respondents rated the departments according to the first criteria as "distinguished," "strong," "good," "adequate," "marginal," and "not sufficient." They were also asked to limit the number of "distinguished" ratings to five. On the other hand, according to the second criteria, the choices of ratings were "extremely attractive," "attractive," "acceptable," and "not attractive."

The 1970 Roose-Andersen study replicated Cartter's (1966) study as the five-year follow-up study. The authors confirmed that the purpose of the study was only informational. They extended the sample from 29 to 36 fields and from 106 to 131 institutions. In all, 2,626 graduate programs were surveyed (Roose and Andersen, 1970).

Both ACE studies were criticized as reflecting three major weaknesses (Lawrence and Green, 1980). The first weakness was specified as the lack of consensus on the meaning of quality. The definition of quality, the authors stated, seemed to vary "from rater to rater, from program to program, and from discipline to discipline, making it almost impossible to compare programs and institutions or to develop normative standards" (p. 8). The second weakness was the halo effect, in which the raters who knew very little about the specific department at an institution might rate it according to their perceptions of the

prestige of the institution as a whole (Astin and Solmon, 1981; Dent, 1978; Elton and Rodgers, 1973; Koelsch, 1981). The alumni effect was identified as another weakness. This was the tendency of raters to overrate their alma maters. It complicated the situation in that the institutions producing the largest number of doctorates also produced the largest number of raters (Lawrence and Green, 1980).

Some scholars used departmental productivity as an important criterion in determining departmental excellence, while many of them suggested its weaknesses. Smith and Fiedler (1971) noted three ways to assess departmental excellence: (1) counting a scholar's publication; (2) rating the quality of a scholar's published research; and (3) counting the number of citations made of a scholar's work. Washburn (1980) found that faculty at Stanford University identified the average number of journal articles published per professor per year as the most important criterion in assessing departmental excellence.

Among the three criteria reported by Smith and Fiedler (1971), counting the number of citations was identified as the most acceptable. Citation rates indicated "the impact of an author's publications upon professional colleagues" (Drew and Karpf, 1981, p. 307). The use of rating the quality of a scholar's published research was intuitively sound, but difficult to measure. The use of publication rates in some studies, however, was criticized as being inappropriate to measure the quality of the department for two reasons. First, those researchers counted publication rates only from the highly referenced journals. Scholars who preferred writing books or articles in other journals were not included (Baird, 1980; Dent, 1978; Drew and Karpf, 1981). Second, by emphasizing publication productivity, excellence in

teaching areas of some faculty members, especially at the undergraduate level, was ignored (Drew and Karpf, 1981). DeSouza, Vogeler, and Foust (1981) commented that:

The primary responsibilities of faculty members in undergraduate departments are teaching and advisement. Those instructors who strive for excellence in teaching should not be dismissed as unproductive because they do not measure up to the research productivity. . . . (p. 170).

Those faculty in the graduate departments might teach far fewer classes each semester than the teachers in the undergraduate departments.

Webster (1981a) did not totally agree with Smith and Fiedler in the use of citation counting. The author believed that counting citations had some obvious virtues. First, they were useful in assessing the importance and influence if one agreed that the often cited articles were more influential and important than the seldom cited ones. Second, they showed the present or the very recent influence of a department's publications. The professor who had done a significant work 20 years ago but had published very little since was unlikely to have a great many recent citations.

Counting citations had many disadvantages as well. First, they were immensely influenced by the popularity of the field of study. In other words, a scholar who was publishing in a growing area was more likely to be cited than others in a less popular area. Second, they were also influenced by the expansion of the field of study. A scholar whose field of study expanded over time was likely to be cited more often than those whose fields did not expand. Third, the citation indexes did not distinguish in any way among "good," "neutral," or "bad" citations. Lastly, counting citations was easily susceptible

to manipulations. For example, in the social sciences, "people like Freud, Marx, Weber, and Piaget were cited hundreds of times every year; the average scholar was cited far less" (Webster, 1981a, p. 23).

Using the 1975 Social Science Citation Index (SSCI) as a data base, Endler, Ruston, and Roediger (1979) ranked psychology departments using total citation and total publication counts. The study received some criticisms regarding certain inaccuracies. There were problems in properly identifying faculty members in the SSCI due to the same individual being listed under two or three different names. Similarly, there was confusion between psychologists with identical surnames. As a result, these errors considerably altered the rankings of certain individuals as well as entire departments (Rossenberg, 1979; Swartz, 1979).

Two years later, Bridgwater and Walkenbach (1981), focusing on the Endler, Ruston, and Roediger (1979) study, conducted a study to determine the accuracy of publication counts obtained from the SSCI in comparison to self-reported publications. Through another research project, the researchers had access to self-reported publications from 1965-1979 for 298 psychologists. During these years, publication counts were broken down into 16 categories, including journal articles, paper presentations, books, book chapters, monographs, etc. The researchers also refined their data pool to include only subjects who listed journal publications in 1975. Moreover, to avoid problems that occurred in the Endler, Ruston, and Roediger study, they only used subjects with unambiguous names. As a result, the subject pool included 113 psychologists whose 1975 journal publications ranged from one to seven. The total of 260 journal articles published in 1975 were reported.

The results of the three analyses showed inaccuracies of publication counts. First, following Endler, Ruston, and Roediger's (1979) exact methodology, the 1975 SSCI Source Index was used to estimate what their count would have been for each of 113 psychologists. Only 174 of the 260 reported publications could be found in the 1975 Source Index. Correlation between an individual's SSCI count and self-reported publications was only .61. It was "surprisingly low given the extensive number of journals (over 3,000) referenced in the 1975 SSCI" (Bridgwater and Walkenbach, 1981, p. 2). Second, the researchers correlated the 1975 publications listed in Psychological Abstracts with subjects' self-reports for that year. The result showed a moderate correlation of .65. Third, while the number of total articles found in the Psychological Abstracts was nearly identical to the number listed in the SSCI, there was only a 35.4% overlap between the articles listed by both sources. Thus, the researchers correlated self-reported publications with the number of articles listed in both sources, eliminating duplicate publications. The researchers reported:

The resulting correlation was increased to .73, but was still a far cry from what would ideally be expected. If Psychological Abstracts or the SSCI were exhaustive and accurate indices of psychologists' journal publications, one would expect correlations in the .90's. That our three analyses yielded much lower correlations questions the accuracy of both archival sources (p. 3).

Not only pointing out the major weaknesses of using publication rates as a measure of departmental excellence, but Drew and Karpf (1981) also suggested that: (1) effective measures of teaching quality be invented; (2) unitary rankings be abandoned; and (3) multidimensional indicators be made. Emphasizing on the last suggestion, Clark, Hartnett, and Baird (1976) found some interesting data in their

pilot study of doctoral program quality. To explore ways to assess departmental quality, a sample of 73 departments were surveyed (24 in psychology, 24 in chemistry, and 25 in history). Four major findings were as follows:

1. Dependable and useful information on educational quality could be obtained at reasonable cost and convenience.
2. Between 25 and 30 measures were identified as especially promising.
3. These measures seemed to be generally applicable across diverse fields.
4. Two clusters of measures emerged: research-oriented indicator (size of department, reputation, physical and financial resources, student academic ability, and faculty publications); and educational experience indicators (educational process, academic climate, faculty interpersonal relations, and alumni ratings of dissertation experiences).

In the spring semester of 1979, 233 faculty members of a large midwestern university completed two questionnaires: the Departmental Evaluation Survey (DES) and the Administration Evaluation Survey (AES). These faculty members represented six departments from a variety of disciplines. The return rate ranged from 70 to 94 percent (Hengstler, Bradenburg, Braskamp, and Smock, 1981). A factor analysis of the DES revealed three underlying dimensions: Satisfaction with the department's Academic Environment, Satisfaction With the Department's Governance and Operating Procedures, and Satisfaction With Student Quality. The same analysis was performed on the faculty ratings of the AES using identical procedures. The results indicated a

one-factor solution which was labeled "Overall Impression of the Administrator's Performance" (p. 262). Additionally, both questionnaires were found to be highly reliable and effective in discriminating among departments. "Very high positive correlations were also found between selected DES and AES items, suggesting that the performance of the head plays a very important role in faculty perceptions of the departments, and vice versa" (p. 259). However, according to the researchers, the unidimensionality of the AES might be a function of both generosity error and a halo effect. Faculty members might have a vested interest in the department and its head and give high ratings to the department and its administrator, reflecting a potential generosity error. In a similar manner, faculty members might rate the administrators in terms of overall impression without differentiating specific aspects of their behaviors.

Observing that the best-known studies of the quality of the graduate departments were centered at the doctoral level, Webster (1981b) suggested a way to assess quality of the master's degree programs. The author proposed five criteria:

1. Faculty's scholarship (research abilities of the faculty rated by other scholars in the same discipline).
2. Amount of time faculty devoted to teaching, advising, guiding student research, and student contact in general.
3. Quality of students (measured by GRE aptitude and achievement scores).
4. Student participation in departmental and campus life (measured by the percentage of full-time students, of those who served as graduate assistants, of those who lived on or near campus, and the

absolute number of students in a department who were fully involved in departmental and campus life).

5. Placement record of graduates (initial academic appointments and how well the program served to advance all its students' careers).

Lawrence and Green (1980) also agreed that assessment of quality of master's degree programs had been ignored. Considering the number of master's degrees awarded in 1978 and 1980, "this lack of information on the quality of the degree seems a decided embarrassment" (Lawrence and Green, 1980, p. 14). The authors explained that some observers regarded the receipt of the master's degree as a step toward the doctoral degree. Thus, the quality of the doctoral program should reflect the quality of the master's program. In addition, the master's degree often served as a screen test for students who desired to pursue doctoral degrees. Ones who did not achieve usually were "gracefully eased out" (Lawrence and Green, 1980, p. 14).

In conclusion, Solmon and Astin (1981) summarized that earlier rating studies of graduate education had employed similar methodologies. The unit of analysis was the department rather than the institution. The most widely used criterion to determine departmental excellence was faculty eminence through scholarship; faculty achieved such eminence through "publication, and peer review of publication underpins the academic reward system" (Solmon and Astin, 1981, p. 23).

Criteria for Determining Departmental Excellence: Undergraduate Level

The assessments of departmental excellence were focused on both graduate and undergraduate education, but "those for graduate

departments have employed the most systematic procedures and generated the most interest among academicians" (Solmon and Astin, 1981, p. 23). The fewer attempts at the undergraduate departments, stated Lawrence and Green (1980, p. 23), were not surprising due to "the scope, diversity, and multiplicity of functions that characterize the undergraduate domain." These characteristics made meaningful comparisons among departments difficult.

Most of the earlier studies on the quality assessment of the undergraduate education centered at rating the institution as a whole. None of them contributed to the academic departments. Criteria used were institutional resources (Jordan, 1963; Brown, 1967; Astin, 1977), student achievement (Krause and Krause, 1970, Dubé, 1974; Tidball and Kistiakowski, 1976; Astin and Hensen, 1977).

Not until the Gourman ratings, which started in 1967 and continued in 1977, had the systematic approach in assessing quality of the institution of higher education focused on departmental excellence and defined what excellence meant. Beginning in 1967, 1,187 four-year colleges were scored on two criteria: (1) strength of the institution's academic departments and (2) quality of nondepartmental areas. Colleges were graded according to the College Board Scale: A=800, B=600, C=400, and D=200. Variable scores in each criteria were averaged to produce a numerical "average academic departmental rating," "average nondepartmental rating," and an overall "Gourman rating" for each institution (Gourman, 1967). The 1977 ratings used a format identical to that of the 1970 Roose-Andersen study. Sixty-eight undergraduate programs were ranked, as well as pre-medical and pre-law

programs in the United States and foreign universities and professional schools (Gourman, 1977).

Both Gourman ratings received considerably strong criticisms. Gourman rated older college faculty members more highly than younger ones on the basis that "a minimum of ten years after college graduation is necessary to produce an excellent teacher in the classroom" (as cited by Lawrence and Green, 1980, p. 36). Further, equal weight was given to ratings of a college's alumni association, faculty effectiveness, public relations, library, and athletic academic balance, "even though common sense suggests that these factors differ considerably in the magnitude of their contributions to institutional quality" (Lawrence and Green, 1980, p. 36). Webster (1981b) commented that Gourman had bias toward large institutions in that he tended to rate large public institutions higher than small liberal arts colleges.

In contrast to the Gourman ratings, Solmon and Astin (1981) explored the feasibility of rating undergraduate departments in a controlled, consistent manner. This study was the Higher Education Research Institute pilot study in which the intent was to "quantify the subjective judgements of undergraduate departmental quality on multiple dimensions" (Solmon and Astin, 1981, p. 24). Approximately 15,000 raters, who were department members in seven fields at all four-year colleges and universities in California, Illinois, North Carolina, and New York, were participants in this study. The seven fields included were: biology, business, chemistry, economics, English, history, and sociology. Using a five-point scale (outstanding, superior, average, marginal, poor), respondents were asked to rate undergraduate departments on six criteria:

1. Overall quality of undergraduate education.
2. Preparation of students for graduate or professional school.
3. Preparation of students for employment after college.
4. Faculty commitment to undergraduate teaching.
5. Scholarly or professional accomplishments of faculty.
6. Innovativeness of curriculum and pedagogy.

Essentially, the findings indicated that the raters basically used two criteria in determining excellence of the undergraduate departments: (1) scholarly excellence of faculty and (2) commitment in teaching. The halo effect was reported to exist among the undergraduate schools in business. The overall reputation of an institution or its graduate program might lead the raters to attribute high quality to underserving departments. For example, Princeton ranked in the top ten departments of business on each of the six criteria, although it did not offer an undergraduate business major. In the later issue of the same journal, Astin and Solmon (1981) added that the halo effect was also in operation when the ratings of one department might suffice as an estimate of the quality in other departments of the institution. Lastly, there also existed the rater bias in this study. The authors stated: "In every instance the biases were in the direction of favoring institutions in one's own home state" (p. 18).

Comprehensive Colleges and Universities

Origin

Comprehensive universities and colleges have emerged as new, typically American institutions in response to the social forces which swept the United States after World War II. Specifically they are the institutional

expression of the synthesis which resulted when an elitist academic emphasis on thought, research, and specialization clashed with social demands for equality, action, accountability, and service (Smith, 1978, p. 469).

Comprehensive colleges and universities were a product of the post-World War II years in which the commitment to social justice was revived by the effort of egalitarianism. The role of higher education in this commitment was described as the emphasis on mass higher education and meaningful access as a means to the end; and equality of opportunity to relevant education dominated the development of post-secondary education for the next three decades (Levine, 1978).

The origin of comprehensive institutions of higher education was spurred by two very important phenomena: (1) the launching of Sputnik by the Soviet Union in 1957 and (2) the Supreme Court decision in *Brown vs. the Board of Education* (Smith, 1978). First, the launching of the Sputnik had an effect on the changes in American higher education in terms of the increase of educational level and complexity of society which stimulated both elitism and egalitarianism. Through the G.I. Bill, the advocacy of the President's Commission on Higher Education for Democracy, and the 1958 National Defense Education Act, college enrollments grew from 18 percent in 1941 to 40 percent in 1964. According to Trow (1974):

What these numbers conceal are two fundamentally different processes. One of these is the expansion of elite universities. . . . The other is the transformation of elite university systems into systems of mass higher education performing a great variety of new functions (at least new to universities) for a much larger proportion of the university age group (p. 61).

Second, the case of *Brown vs. the Board of Education* reflected the egalitarian effort on the "development of the technology of protest by

the Black movement, and the adoption of successive affirmative action/ equal opportunity acts covering race, age, ethnic origin, and a wide variety of handicaps" (Smith, p. 471).

During these years of educational development, the research university and the liberal arts college grew toward their pure form--the form of intellectual elitism. As a result, they became less capable of dealing with the demand of mass education. The gap between the professional education on the one hand and the liberal arts education on the other was filled by the emerging comprehensive colleges and universities since that time (Centra, 1977).

Concept of Comprehensive Universities and Colleges

The attempts to organize and classify the higher education institutions had long been a concern of the Carnegie Commission on Higher Education. In 1970, the Commission sponsored Harold Hodgkinson to conduct the study on the profile of change in higher education. In this statistical report, Hodgkinson (1971) classified the institutions of higher education by using four criteria: (1) type of control (public, sectarian, private, and nonsectarian); (2) institutional size (small, medium, large, giant, and super); (3) geographic area (New England, Middle Atlantic, Southeast, Great Lakes, Plain, Southwest, Rocky Mountain, and Far West); and (4) comprehensiveness (two-year institutions, four-year institutions, institutions granting master's degrees, and institutions granting doctoral degrees). The comprehensive colleges and universities did not yet exist.

The term "Comprehensive Universities and Colleges" was first used in the Carnegie Commission's classification of institutions in 1973. This classification was based on 1970 data and was the result of several years of work by the Commission. However, the use of the term "comprehensive" no longer stood in the literature of higher education. Rather, it was frequently used by the large public research institutions which were more comprehensive in the common use of the word.

In 1976, the Carnegie Council on Policy Studies in Higher Education, the successor to the Commission, published the revised edition of the classification of the institutions. In this new classification, the Comprehensive Universities and Colleges I were defined as the group of institutions that offered a liberal arts program as well as at least two professional programs such as engineering and business administration. The required enrollments were a minimum of 2,000. The Comprehensive Universities and Colleges II, on the other hand, were those that offered a liberal arts program and at least one professional program. Enrollments were at least 1,000 for public institutions and at least 1,500 for private institutions (Carnegie Council on Policy Studies in Higher Education, 1976; Levine, 1978).

Among the limited literature on the comprehensive colleges and universities was Smith's (1978) proposal of a new way of defining this type of institution. He called them "synthesizers of liberal education and professional education," because, as once stated, they filled the gap between the research university and the liberal arts university which specialized in professional education and liberal education, respectively (p. 469). They were, in essence, the balance between theory and application. The potential balance between theory

and application had two dimensions, according to Harclerod, Sagen, and Molen (1969):

One dimension is composed of college versus university status: with emphasis on undergraduate and/or beginning graduate instruction with little attention to research versus an emphasis upon advanced graduate and professional instruction with additional major attention to research and, in some institutions, to high-level public service. The second dimension is applied or practical orientation versus a theoretical orientation. Institutions having the former characteristics tend to emphasize preparation for terminal occupations and applied research, and service involving solutions to immediate problems. Institutions characterized by a theoretical frame of reference tend to emphasize the basic liberal arts subjects and basic research with less attention to practical application (p. 113).

The authors further stated that the institutions occupying the middle arena could be characterized as comprehensive because of their emphasis on both theory and application. However, "few can maintain such a balance" (p. 113).

Harclerod et al. (1969) also observed that the factors that moved the institutions toward advanced instruction and research and toward a theoretical rather than applied orientation were the internal dynamics of institutional development driven by the aspirations of professionalized faculty. This resulted in the rapid expansion of doctoral programs, the emphasis on research, and the solidification of the role of discipline and departments in the structure and governance of the institutions. Research universities and traditional liberal arts colleges emphasized these values.

Smith (1978) also believed that the reason that comprehensive colleges and universities are capable of maintaining the balance between theory and practice was the institutional funding. The funding of comprehensive colleges and universities was basically enrollment

driven. For this reason, they were very responsive to social needs and student interests which acted as "a counterweight to the academic interests of the faculty" (Smith, 1978, p. 473). The dichotomy between liberal and professional education, according to Smith, only existed in the respect that the purpose of the liberal arts major was problem solving for further understanding and knowledge, while the purpose of the professional major was to join the understanding and knowledge with the real world to create an activity to solve a problem. Additionally, Smith also realized the tension between theory and application, but he strongly believed that this dynamic balance could only coexist within the comprehensive colleges and universities, where "the areas of similarity can be emphasized while the essential differences can be maintained" (p. 481).

Another proposed idea that Smith (1978) had was the inclusion of Liberal Arts Colleges II (the term used by the Carnegie Council) in the comprehensive concept. Criticizing that the Carnegie's classification was based solely on the size of organization and inclusiveness, the author commented that the restriction "obscures an important characteristic of American higher education" (p. 482). He further noted that the colleges characterized as Liberal Arts Colleges II were much closer to the comprehensive concept than they were to those of selective liberal arts colleges. A substantial number of Liberal Arts II awarded more than 50 percent of their degrees in occupational programs. This was an indication of a "giving away" tradition of liberal arts colleges. These colleges were, in fact, "mini-comprehensives" (Smith, 1978, p. 482).

The only obstacle in Smith's (1978) idea of including Liberal Arts Colleges II with the Comprehensive Universities and Colleges I and II was the identity crisis of these liberal arts colleges. They considered themselves liberal arts colleges, and had been considered as such by others (Astin and Lee, 1972). Astin and Lee called these institutions the "invisible" colleges. They also emphasized that "at the heart of this dilemma is the confusion of the invisible college over its role and identity" (p. 95). Moreover, Guardo (1978) also agreed with Astin and Lee as she stated:

In recent years, many Liberal Arts II Colleges have added professional or career oriented programs to their curricula in an attempt to counteract adverse enrollment trends. Hence, within a single college entity liberal and professional programs may coexist; that they interact to attain a comprehensive character, however, remains problematic (p. 489).

Guardo also found that many of these colleges still perceived themselves as having educational missions comparable to those of Liberal Arts I colleges.

Recent Studies

With the same criticism Smith (1978) had against the Carnegie classification, Korb (1982) classified colleges and universities using the empirically determined system. Essentially, the researcher used the following dimensions to cluster institutions: tuition dependence, instructional emphasis, black student enrollment, facilities, research emphasis, growth, size, and endowment level. The institutions were categorized as major doctoral/research institutions, major doctoral/non-research institutions, comprehensive institutions, and general

baccalaureate institutions. The results of the study indicated that there were 287 institutions under the comprehensive institution category.

Lane, Stenlund, and Westlund (1982), Swedish educators who conducted the first national survey of Swedish teachers and researchers, revealed that the general properties of the institutions characterized as comprehensive were: (1) integration of organizational units, (2) deconcentration of higher education resources, (3) egalitarian values, (4) emphasis on teaching and vocational training, and (5) reorientation of the research system, including an emphasis on interdisciplinary research.

Conclusion

The concept of comprehensive colleges and universities either described by the Carnegie Council (1976) or by Smith (1978) involved the inclusion of liberal and professional education in one entity. The idea that these two extreme areas could be balanced within the context of comprehensive colleges and universities had not yet become reality because, as Harclerod et al. (1969) stated earlier, very few of these institutions were able to maintain this balance. In addition, Smith (p. 483) concluded that "there is no one balance point which is right for all institutions at the same time, nor right for one institution through all time."

Smith's (1978) proposition to include Liberal Arts Colleges II in the arena of comprehensive colleges and universities included one problem--the reluctance of liberal arts colleges to assume a new identity. It seemed more important to them to maintain their identity

as liberal arts colleges because of prestige and esteem they had earned than to become a comprehensive college or university. Another example could be illustrated by Smith's (1978) experience at Drake University. Not knowing the ultimate goals of the comprehensive colleges and universities, one of the deans examined the list of comprehensive universities and said, "Who wants to be one of those?" (Smith, 1978, p. 476).

Criticisms were made that the Carnegie classification was limited to organizational size and inclusiveness of subject matter coverage (Smith, 1978) and separated institutions into very broad, summary categories that in many aspects has as much diversity within classes as between (Korb, 1982). Yet there is not, at present, any better model to organize the institutions of higher education in a systematic way. Smith's model was problematic, while Korb's model had some major limitations. First, Korb's model was only limited to four-year institutions. Second, the only prominent indication of the results was the number of institutions clustering around those criteria. It did not show exactly who were what, as the Carnegie classification clearly did. Smith, however, admitted that the Carnegie classification had been a valuable tool in the study of American higher education.

Summary

Chapter II began with an overview of the academic department, the academic department head, previous studies on criteria used to determine departmental excellence, and finally, examined comprehensive colleges and universities. The academic department, in the first section of this chapter, was identified as the basic administrative

unit of a college or university. It was the vital part for planning and management of the learning process and research, creative activities, and public services. The academic department head, in the second section, was the one who chaired the academic department and made the institution perform its best. He or she was the person who provided intellectual and professional leadership for his or her departmental staff members. The role of the department head had been described as ill-defined, pressuring, ambiguous, and conflicting. In earlier studies, critics tended to agree that most department heads were traditionalists. However, there was no valid proof that their perceptions were outdated. Department heads might not be the most distinguished scholars in the departments, but they were, overall, no less capable and knowledgeable, academically and professionally, than any other members of the department.

In the third and fourth sections, the researcher reviewed previous studies that attempted to measure the excellence of institutions using the academic department as a unit of analysis. The review of related literature in this section reflected and confirmed the idea of the academic department as the most essential part of the administration of the institution. Departmental excellence, in turn, indicated institutional excellence. In addition to the criteria used in determining departmental excellence, strengths and weaknesses of several approaches were identified.

In the final section, the review of literature was limited because of a lack of studies in this area. The basic property of the comprehensive colleges and universities, regardless of the enrollment

size and the inclusiveness of the subject matter, was the combination of professional and liberal education in the same entity.

In conclusion, it was evident that criteria used in determining departmental excellence had not yet been sufficiently developed. They were different from one level of education to another (undergraduate versus graduate education), from one discipline to another in the same institution, and from one institution to another of different types (i.e., research institution versus comprehensive institution). Criteria developed for one type of institution could not be appropriately used with other types of institutions because of differences in institutional mission. In such type of institutions as comprehensive colleges and universities, no attempt in assessing departmental excellence had been conducted before. It was the purpose of this study to identify and analyze the criteria used by department heads at the public comprehensive colleges and universities for determining departmental excellence.

CHAPTER III

METHODOLOGY

In this chapter the researcher outlined the major components pertaining to the methodology that were used in the research study. These components were: (1) design of the study, (2) description of the population and the sample, (3) description of the survey instrument, (4) procedures, including data collection, and (5) data analysis.

Design of the Study

The appropriate type of research for this study appeared to be the descriptive one. A descriptive study determines and reports the way things are (Gay, 1981). This method is useful to gather practical information which may be relevant for the improvement or justification of an existing situation, and can provide a foundation upon which further research can be conducted (Van Dalen, 1966).

Description of the Population and Sample

The population from which subjects were drawn was department heads at the public comprehensive colleges and universities, found in the Carnegie classification list. The subjects for this research study included selected department heads of randomly selected public comprehensive colleges and universities.

Description of the Survey Instrument

Washburn's (1980) multidimensional questionnaire was used in a slightly modified format (Appendix C). Two major justifications for using this specific questionnaire were: (1) the need for consistency in studying departmental excellence and (2) the distinguished key components of the questionnaire (community and university service, faculty scholarship and research, graduate program, students and undergraduate teaching). Thirty-eight criteria of departmental excellence were listed under these five key components. In addition, department heads might propose two additional criteria in each key area. Five short answer questions were included to obtain information concerning the status of the department, the improvement of departmental excellence, and the improvement of institutional recognition as a whole.

Washburn (1980) did not report validity and reliability of his instrument. Therefore, a pilot study was conducted to fulfill this lack of information, as well as to pretest the receptivity of the subjects.

The modification of the original questionnaire involved a change in wording. On page four of the original questionnaire, the phrase "master's degrees" was used to replace "Ph.D.'s," "doctoral," and "Ph.D.'s" in items 24, 25, and 28, respectively. Additionally, the phrase "and receive the Ph.D." was left out from item 25. The reason for the change in language for these three items was that, typically, comprehensive colleges and universities did not offer doctoral programs or award Ph.D.'s; moreover, items 51 and 55 could not be

included in this new questionnaire because item 51 can only be used at Stanford University, while item 55 can only be used with faculty members. As a result, items 52, 53, and 54 will become items 51, 52, and 53, respectively.

Procedures, Including Data Collection

1. Selection of Subjects - Fifty-three public comprehensive colleges and universities were randomly selected from the population of 354 public comprehensive colleges and universities listed by the Carnegie Council on Policy Studies in Higher Education. Based on the Biglan model, department heads, chairmen, or chairpersons within each of these randomly selected institutions included those from eight departments: mathematics, botany, English, psychology, computer science, agricultural economics, economics, and secondary education. The selection was made on the justification that these departments were most likely to exist in these randomly selected institutions. Substitution could be made among departments in the same discipline. Names, addresses, and telephone numbers of the subjects were obtained from current editions of the institution's catalog, available through microfiche at the library of Oklahoma State University.

2. Preparation of Questionnaire - A modified questionnaire was prepared for a pilot study. This pilot study was conducted at Oklahoma State University, a Research University II, as classified by the Carnegie Council. Ten percent of the total graduate and undergraduate department heads were the subjects for the pilot study.

Since there is no way to quantitatively express validity of the instrument, it can be assumed valid by expert judgments (Gay, 1981).

Washburn (1980) reported that the 38 criteria had been previously investigated by some researchers. He had used, reused, and modified the questionnaire several times. Therefore, this questionnaire might be assumed to have validity and was appropriate for this research study after slight modification.

Reliability was estimated by using the test-retest method. The questionnaire was personally delivered to and collected from each subject. One week later, the administration of the same questionnaire was repeated to the same subjects. Then, the proportions of the agreement and disagreement of the responses on the two questionnaires of each subject were determined in terms of percentages. Percentages of the agreement of the responses across all subjects indicated the coefficient of reliability for the questionnaire.

3. Data Collection - The questionnaire consisted of: (a) a cover letter explaining the purpose, procedures, deadline date, and assuring confidentiality (Appendix B), (b) a coded (for identifying nonrespondents) copy of the questionnaire for each department head, and (c) a stamped, self-addressed envelope for each department head to return his or her response. Questionnaires were mailed to selected subjects. A three week period was allowed for the returned responses, at which time reminder postcards were forwarded to the nonrespondents (Appendix D). Overall, 10 full weeks were involved in the data collection process.

Data Analysis

To answer the first research question presented in Chapter I, data analysis was reported in narrative and tabular form. Percentages

of respondents who ranked each key component as first, second, third, fourth, and fifth were calculated. The key component that received the highest percentage in the first category was selected as the most important key component. To rank the rest of the key components, the selection remained the same.

The second research question was also reported in narrative and tabular form. The number of department heads reporting criteria of departmental excellence in terms of "very important," "moderate," and "unimportant" were calculated in terms of percentages. The highest percentage of the criterion reported was interpreted as the most important criterion for determining departmental excellence as perceived by department heads.

To answer the third research question, descriptive analysis was utilized. Procedures were as follows:

1. Department heads were categorized into eight groups as in Biglan's (1973) three dimensional model, based on their departmental affiliations.
2. In each discipline, the percentages of key components were calculated and the highest percentage in each category was selected in the same manner as in the first research question. The comparison among discipline then proceeded.
3. Only the top five most important criteria identified in the second research question were used to compare with the criteria identified by department heads of each discipline.
4. Percentages of responses on each criterion in each discipline were reported in narrative and tabular form.

Lastly, the fourth research question was answered by computing the average of the optimal and minimum acceptable levels of performance of the criteria that entered the top five list.

CHAPTER IV

PRESENTATION AND ANALYSIS OF DATA

In this chapter the researcher presented the finding of the pilot study in estimating the reliability of the instrument, response rate, ranking of the key components, criteria of departmental excellence, the similarities and differences of the key component ranking, criteria of departmental excellence in each discipline, and the levels of acceptability of the five most important criteria. The results of the data analysis were presented in both narrative and tabular forms.

Pilot Study

Due to the lack of information on the reliability of Washburn's (1980) questionnaire, the pilot study was conducted at Oklahoma State University. Eight departments were randomly selected. The questionnaire was personally delivered to and collected from each department head. One week later, the retest session was conducted by mail.

The completion of the questionnaire required the respondents to perform two tasks: ranking the key components and identifying the criteria they considered as "very important," "moderately important," and "unimportant." The reliability coefficients, then, were estimated from the percentages of agreement and disagreement of the responses on the two questionnaires of each subject performing each task.

First, when the rankings of key components on both questionnaires of each subject were compared, it was found that the percentage of agreement of the responses of each subject was 100 percent. In other words, the department heads were very consistent in prioritizing the key components. Six of the eight department heads ranked "faculty scholarship and research" first; "graduate program" second; and "students," "undergraduate teaching," and "community and university service" as third, fourth, and fifth, respectively. The other two department heads differed in their opinions on the first three key components. "Faculty scholarship and research" was ranked first by one department head, but was ranked second by the other. Graduate program component was ranked second by the former, but was ranked third by the latter. Finally, undergraduate teaching component was ranked third by the former, but was ranked first by the latter. The fourth and fifth components were "students" and "community and university service," respectively. The rankings of key components are reported in Table I.

It was evident in the ranking of key components from first to fifth (first being most important and fifth being least important) that faculty scholarship and research was ranked first, the graduate program was second, student component was third, and undergraduate teaching and community and university service ranked fourth and fifth, respectively. Washburn (1980) reported exactly the same ranking pattern. Percentages of responses from first to fifth in the Washburn study were 91, 66, 44, 49, and 69.

Second, by comparing the listing of criteria for departmental excellence as "very important," "moderate," and "unimportant" in both

responses of each department head, the means of percentages of agreement were 77.9%, 82.0%, 89.3%, 92.0%, 63.0%, 84.6%, 79.9%, and 81.0%. Weighted (or pooled) mean was 81.2%. This mean value represented the second reliability coefficient of this instrument.

TABLE I
RANKING OF KEY COMPONENTS (PILOT STUDY)

Component/Rank	First (%)	Second (%)	Third (%)	Fourth (%)	Fifth (%)
Community and University Service	--	--	--	--	100
Faculty Scholarship and Research	88	12	--	--	--
Graduate Program	--	88	12	--	--
Students	--	--	75	25	--
Undergraduate Teaching	12	--	12	75	--

Response Rate

Fifty-three public comprehensive universities and colleges were randomly selected from the population of 354 public universities and colleges listed by the Carnegie Council on Policy Studies in Higher Education (1976). Four hundred questionnaires were mailed to the

subjects. They included primarily the department heads, chairmen, or chairpersons from eight departments: mathematics, botany, English, psychology, computer science, agricultural economics, economics, and secondary education. Substitutions of chemistry to mathematics, microbiology or zoology to botany, mechanical engineering to computer science, and special education and educational administration and supervision to secondary education were made in some institutions. Agriculture, though, was not classified by Biglan (1973) under applied-hard life system discipline, and was used to substitute agricultural economics in some institutions because there were no agronomy, dairy science, and horticulture departments in these institutions. Twenty-two of the 53 randomly selected institutions had substitutions made for departments within the category of Biglan's taxonomy. The number of times particular departments used substitutes within disciplinary areas is shown in Table II. A total of 32 of 400 departments (14%) were substituted by other departments within their disciplinary areas. Under the dissertation assumption that department heads within the same disciplinary area of Biglan's taxonomy used the same criteria in determining departmental excellence, and with the fact that a small percentage of the total number of departments used substitution, it was determined that the process should not affect the validity of the study.

The response rate at due date was 58% (232 questionnaires). Eighteen of these questionnaires were unusable. As a result, the total usable questionnaires was 214 (53.5%). Follow-up activities were conducted, and 60 more questionnaires were returned. The total returned and usable questionnaires were 274 (73%). The questionnaires

were classified into eight disciplines according to Biglan's model, and the total number of responses in each discipline is shown in Table III.

TABLE II
SUBSTITUTIONS OF DEPARTMENTS WITHIN
DISCIPLINARY AREAS

Departments Selected Within Disciplinary Area	No. of Times Actual Departments Were Used	No. of Times Sub- stitutions Occurred
Agricultural Economics	23	7
Botany	38	1
Computer Science	44	6
Mathematics	41	9
Secondary and Continuing Education	41	9
Totals	187	32

Ranking of the Key Components

Table IV shows the rankings of relative importance of the five key components of departmental excellence used on the questionnaire. More than 5 out of 10 respondents ranked "faculty scholarship and research" as the most important component for judging departmental

excellence, and ranked "community and university service" as the least important component. "Graduate program" was ranked second and also third, with percentages of 35 and 37, respectively. This ranking pattern of the first, second, and fifth corresponded to Washburn's (1980) results in which the author stated that the results were not surprising because these three components constituted the "three traditional categories into which departmental functions have been classified for at least the last twenty-five years" (p. 146). "Students" was ranked as the fourth most important component among five key components.

TABLE III
NUMBER OF RESPONSES IN EACH DISCIPLINE

	Hard		Soft	
	Nonlife	Life	Nonlife	Life
Pure	N=37	N=34	N=35	N=42
Applied	N=29	N=23	N=33	N=41

Criteria of Departmental Excellence

After ranking five key components, respondents were asked to review a list of 38 possible characteristics of departmental excellence.

Respondents circled the characteristics they considered very important and crossed out the ones they considered unimportant. The questionnaire provided no limitations on the minimum or maximum number of characteristics to circle or cross out (Appendix C).

TABLE IV
RANKING OF KEY COMPONENTS

Component/Rank	First (%)	Second (%)	Third (%)	Fourth (%)	Fifth (%)
Community and University Service	--	10	14	22	53
Faculty Scholarship and Research	55	16	14	12	2
Graduate Program	4	35	39	14	8
Students	8	27	22	31	12
Undergraduate Teaching	39	20	16	18	4
Totals*	106	108	103	97	79

*The ranks do not sum to 100% because respondents had the option to give two or more categories the same rank.

On the average, department heads most often circled characteristic B11 (average number of refereed journal articles published per professor per year). Eighty-four percent of the respondents circled

this characteristic, while only one respondent marked it as unimportant. The second most important characteristic regarded by the respondents was characteristic E44 (average overall student evaluation of departmental courses). Characteristic B15 (proportion of faculty conducting research or scholarly work with external funding) was regarded as the third most important characteristic for judging departmental excellence. The fourth and fifth characteristics were listed under the "community and university service" component. They were: A3 (proportion of departmental faculty serving on a national review board or professional advisory committee), and A1 (proportion of departmental faculty serving on a university committee), respectively. Overall, the top five departmental characteristics were regarded as very important by at least 55% of the respondents (Table V).

While B11 was chosen as the most important criterion in determining departmental excellence, A5 (proportion of faculty who serve as general advisers for undergraduates without a major) was at the bottom of the "very important" list (8%). However, it was at the top of the "unimportant" list while B11 was at the bottom of this list. This pattern of responses corresponded with the ranking pattern in which B (faculty scholarship and research) was ranked as the most important component, and A (community and university service) was ranked as least important. It was also found that 64% of the respondents who ranked "faculty scholarship and research" as first also regarded B11 as the most important criterion of departmental excellence.

While the key component C (graduate program), however, was ranked both second and third, and key component D (students) was ranked fourth, none of the departmental characteristics under these two

TABLE V

LISTING IN RANK-ORDER OF THE PERCENTAGES OF
RESPONDENTS DESIGNATING EACH CHARACTER-
ISTIC AS VERY IMPORTANT, MODERATELY
IMPORTANT, AND UNIMPORTANT

Rank	Very Important		Rank	Moderately Important		Rank	Unimportant	
	Char.	%		Char.	%		Char.	%
1	B11	84	1	E46	69	1	A5	59
2	E44	63	2	B9	57	2	D33	53
3	B15	61	3	B10	53	3	B17	45
4	A3	59	4	B8	49		B18	45
5	A1	55		E41	49		C22	45
6	D34	49	6	B13	45		D37	45
7	B14	45		E43	45	7	C24	42
	C25	45	8	B14	43	8	C26	39
	D31	45		B16	43		D36	39
	D35	45		B24	43	10	B16	35
	E42	45		B34	43		C21	35
12	B12	43	12	A2	41		D38	35
	C27	43		C22	41	13	A2	33
14	D32	41		C26	41		C27	33
15	C31	39		D37	41	15	A4	31
	D36	39		E42	41		D32	31
17	A4	37		E45	41		D35	31
	C23	37	18	B18	39	18	C25	29
19	B8	35		C28	39	19	B12	27
	B10	35	20	A1	37		B13	27
	C28	35		B15	37		C23	27
	E43	35		C23	37		C28	27
23	B9	33		D31	37		E45	27
	D38	33	24	A3	35	24	D31	20
	E45	33		A4	35		E41	20
26	B13	31		A5	35	26	B8	18
	E41	31		B12	35		E43	18
28	A2	29		D33	35		E49	18
	B17	29		D38	35	29	B10	14
30	B16	27	30	D32	31		B14	14
31	E46	22	31	B17	29	31	B9	12
32	C26	20		C21	29		E42	12
33	B18	18	33	C25	27	33	A1	10
34	C22	16		D35	27		D34	10
	D37	16		E44	27		E44	10
36	C24	14	36	C27	24	36	A3	8
	D33	14		D34	24	37	B15	4
38	A5	8	38	B11	16	38	B11	2

components appeared in the top five most important criteria. Additionally, while key component A (community and university service) was ranked as the least important component, two departmental characteristics under this component were ranked fourth and fifth in the top five most important criteria for judging departmental excellence.

The instrument was designed to allow respondents to provide additional comments on departmental characteristics that brought departmental excellence. Forty-seven percent of the respondents noted that publication of scholarly research in the refereed journal articles was the most important departmental characteristic that brought departmental excellence. Two additional characteristics were added by the respondents: (1) the ability of the department to obtain external funds (18%), and (2) the university teaching and service (16%).

Similarities and Differences in Criteria Used for Judging Departmental Excellence Among Eight Disciplines

In the third phase of data analysis, department heads were categorized into eight groups as in Biglan's (1973) three-dimensional model, based on their departmental affiliations. The ranking of key components in each of the eight disciplines is shown in Table VI. Six of the eight disciplines ranked "faculty scholarship and research" (B) as the most important key component. The two disciplines differed were "applied-hard life system" and "applied-soft life system" disciplines. Both of them ranked "undergraduate teaching" as the most important key component.

TABLE VI
RANKING OF KEY COMPONENTS IN EIGHT DISCIPLINES

Task Area:	Hard						Soft					
	Nonlife Rank	System Component	%	Life Rank	System Component	%	Nonlife Rank	System Component	%	Life Rank	System Component	%
Pure	1	B	43	1	B	75	1	B	60	1	B	55
	2	E	33	2	D	65	2	E	60	2	C	37
	3	D	32	3	C	50	3	C	80	3	D	36
	4	C	43	4	E	50	4	D	50	4	D	36
	5	A	71	5	A	65	5	A	40	5	A	55
Applied	1	B	60	1	E	75	1	B	80	1	E	49
	2	C	60	2	C	75	2	C	80	2	E	49
	3	D	60	3	B	50	3	E	50	3	C	49
	4	E	30	4	D	50	4	D	40	4	B	49
	5	A	20	5	A	35	5	A	60	5	A	37

Note: A = Community and University Service
 B = Faculty Scholarship and Research
 C = Graduate Program
 D = Students
 E = Undergraduate Teaching

Furthermore, all eight disciplines agreed that "community and university service" (A) was the last among five components, although the percentages of agreement varied in each discipline. "Graduate program" (C) was ranked second by four disciplines and was ranked third by three disciplines. "Students" (D) was ranked third by three disciplines and fourth by four disciplines (Table VI). (Tables IX through XVI in Appendix E show the rankings of key components of each discipline.) In the pure-soft life system discipline, key component D (students) was ranked both third and fourth with 36%. Likewise, key component E (undergraduate teaching) in the applied-soft life system discipline was ranked both first and second with 49%.

As reported earlier, the five most important criteria used in determining departmental excellence were: (1) B11 (average number of refereed journal articles), (2) E44 (average student evaluation of courses), (3) B15 (proportion of faculty conducting research with external funding), (4) A3 (proportion of faculty serving on national review board or professional advisory committee), and (5) A1 (proportion of faculty serving on university committee). These criteria were used as a basis of comparison to determine the similarities and differences among eight disciplines. The top five criteria selected as most important for judging departmental excellence in each discipline are listed in Table VII. In the disciplines other than the pure-hard life system and the applied-hard nonlife system, some ranks were missing because several departmental characteristics received the same percentages.

It was found that five of eight disciplines selected characteristic B11 as first in rank. The percentages of agreement of respondents

TABLE VII
LISTING OF RANK-ORDER (1-5) OF THE PERCENTAGES OF
RESPONDENTS OF EACH DISCIPLINE DESIGNATING ONLY
AS VERY IMPORTANT CRITERIA

Task Area:	Hard						Soft					
	Nonlife System			Life System			Nonlife System			Life System		
	Rank	Criterion	%	Rank	Criterion	%	Rank	Criterion	%	Rank	Criterion	%
Pure	1	A1	86	1	B11	100	1	B11	100	1	B11	91
		A3	86	2	B15	75	2	C25	80	2	E44	73
		B15	86	3	C25	50	3	E44	60	3	E42	55
	4	B11	71	4	C27	45		A1	60		D31	55
	5	E44	57	5	D31	42		D34	60	5	D32	45
		D34	57									
Applied	1	A3	52	1	A3	78	1	B11	100	1	B11	85
	2	E44	48	2	A1	74	2	A3	80	2	B15	73
	3	B11	44	3	B10	69		B15	80		C25	73
	4	B15	41		D34	69		E44	80	4	D35	61
	5	D34	35	5	E42	65	5	B15	60	5	A1	59

Note: A1 = Proportion of departmental faculty serving on a university committee
A3 = Proportion of departmental faculty serving on a national review board or professional advisory committee
B10 = Proportion of faculty elected to a national or professional honor society
B11 = Average number of refereed journal articles published per professor per year
B14 = Proportion of faculty serving as a journal editor or referee.
B15 = Proportion of faculty conducting research or scholarly work with external funding
C25 = Proportion of master's degree students who complete the program
C27 = Proportion of graduate students who publish their work while in the program
D31 = Average SAT Verbal score of undergraduate majors
D32 = Average SAT Mathematics score of undergraduate majors
D34 = Proportion of undergraduate majors who graduate and pursue graduate training in the discipline
D35 = Average GRE Verbal score of entering graduate students
E42 = Proportion of departmental faculty who teach at least one undergraduate course
E44 = Average overall student evaluation of departmental courses

in these disciplines ranged from 85% to 100%. The applied-hard non-life system discipline regarded characteristic B11 as third most important criterion, while the pure-hard nonlife system discipline regarded it as fourth in rank. Only one discipline, applied-hard life system discipline, did not list characteristic B11 in the top five.

Characteristic E44 (student evaluation of courses) was considered the second most important criterion by three disciplines: pure-soft life system, applied-hard nonlife system, and applied-soft nonlife system. It was considered third in rank and fifth in rank by the pure-soft nonlife system discipline, and the pure-hard nonlife system discipline, respectively. Three other disciplines did not list it in their top five.

The pure-hard nonlife system discipline highly regarded characteristic B15 (proportion of faculty conducting research with external funding) as first in rank (86%), together with characteristics A1 (proportion of faculty serving on university committee) and A3 (proportion of faculty serving on national review board or professional advisory committee). Characteristic B15 was second in rank in the pure-hard life system, applied-soft nonlife system, and applied-soft life system disciplines. It was selected as the fourth most important criterion by the applied-hard nonlife system discipline.

Characteristic A3 was selected as the first most important criterion by three disciplines: pure-hard nonlife system, applied-hard nonlife system, and applied-hard life system. Percentages ranged from 52% to 86%. It was second in rank as selected by the applied-soft nonlife system discipline.

Lastly, characteristic A1 was selected as first, second, third, and fifth by the pure-hard nonlife system, applied-hard life system, pure-soft nonlife system, and applied-soft life system disciplines, respectively. The other four disciplines did not list it in their top fives.

Levels of Acceptability

After ranking each of the five key components and marking the relative importance of departmental characteristics recorded under each key component, department heads were asked to specify levels of acceptability of only the characteristics that were considered "very important." Levels of acceptability were two numbers: one specifying the optimal level of performance, and the other specifying the minimum level of performance. Department heads had the option of writing "N.A." if they felt a number or proportion could not be appropriately specified. They also had the option to write a "?" if they had no informed opinion regarding an optimal or a minimum level of performance for a particular characteristic. Because the five most important criteria for determining departmental excellence as identified in Table V would be widely used, only the optimal and minimum levels of acceptability of each of these five criteria were presented here.

The average number of refereed journal articles published per professor per year (characteristic B11) ranged from as low as one to as high as eight. On the average, department heads reported that 2.6 refereed journal articles published per professor per year was the optimal level of performance. Only one respondent felt that specifying level of performance was inappropriate.

Of the top five criteria, characteristic E44 (the average overall student evaluation of departmental courses) was the most distinctive. Scores could range between one (minimum level) to seven (optimal level). Washburn (1980) described that, by tradition, student evaluation of departmental courses ranged from a low of one to a high of seven. In the present study, seven was the number mentioned most; the mean optimal level was 6.5. Eighty percent of the department heads responding on this characteristic reported an optimal average student evaluation of 6.0 or higher. The average minimum level of performance was 2.8.

The specified optimal level of acceptability of the proportion of faculty conducting research with external funds (characteristic B15) ranged from 10% to 80%; whereas, the minimum level ranged from 15% to 30%. Means of optimal level and minimum level were specified at 43.4% and 16.5%, respectively.

The averages of the optimal and minimum acceptable levels in the last two criteria (A3 and A1) were the same. The average optimal level of the proportion of departmental faculty serving on a national review board or professional advisory committee (characteristic A3) and the proportion of departmental faculty serving on a university committee (characteristic A1) was 25%. The average minimum acceptable of each criterion was 10%.

The summary of the optimal and the minimum levels of performance of each of the five criteria is reported in Table VIII. On the five widely regarded criteria, there was no unanimity of the optimal or minimum acceptable levels of performance that predominated across disciplines.

TABLE VIII
THE OPTIMAL AND THE MINIMUM LEVEL OF
PERFORMANCE OF EACH OF THE FIVE
MOST IMPORTANT CRITERIA

Criterion	Mean Optimal Level	Mean Minimum Level
Average number of refereed journal articles published per professor per year	2.6	1.0
Average overall student evaluation of departmental courses	6.5	2.8
Proportion of faculty conducting research with external funds	43.4%	16.5%
Proportion of departmental faculty serving on a national review board or professional advisory committee	25.0%	10.0%
Proportion of departmental faculty serving on university committee	25.0%	10.0%

Summary

Chapter IV began with the presentations of the results from the pilot study, the response rate of the actual research study, the ranking of the key components of the respondents, the selected criteria used in determining departmental excellence, the similarities and differences of the selected criteria in each of the eight

disciplines, and finally, the levels of acceptability of the selected criteria. In the pilot study, the test-retest method in estimating reliability of the instrument was utilized. It was found that the instrument was of acceptable reliability because the averaged percentage of agreement of each subject on responses collected with a one week time interval between collection periods was 81.2%. The result was also confirmed that respondents in the pilot study were consistent in their rankings of key components. The percentage of agreement of the ranking was 100% for each subject.

The overall returned and usable responses were 274, which constituted 73% of the total samples, or roughly 10% of the populations. The ranking of the five key components was reported in the third section of this chapter. The ranking from first to fifth key component was: (1) faculty scholarship and research, (2) and (3) graduate program, (4) students, and (5) community and university service.

The results of the study indicated that the most important criterion in determining departmental excellence was the average number of refereed journal articles published per year (B11). Four other "very important" criteria were identified to compose the top five list. The top five criteria indicated by department heads of each discipline were compared to this list to determine the similarities and differences among disciplines. The findings indicated that the similarities were evident only in the most important criterion. Five out of eight disciplines regarded B11 as the most important criterion, while other criteria differed in terms of percentages of agreement and the ranks that were given. Lastly, the optimal and the minimum acceptable levels of performance on each of the five criteria were specified.

CHAPTER V

FINDINGS, CONCLUSIONS, AND RECOMMENDATIONS

The present study was originated by the fact that there had not yet been an agreement among scholars in higher education as to which criteria constituted excellence of the academic department. With the belief that each type of institution should have its own criteria for judging excellence, the literature review was conducted. It was apparent, as a result, that the assessment of departmental excellence at the comprehensive universities and colleges had been largely ignored. Most studies on institutional excellence focusing on the academic departments were conducted at well-known research universities. This research study was the first attempt to single out the most important criteria that department heads at the public comprehensive universities and colleges used in determining departmental excellence. The five most important criteria were compared to the five criteria that department heads of each discipline indicated as being most important in defining departmental excellence. The criteria of departmental excellence for the eight disciplines were then analyzed for differences and similarities.

With the permission of Washburn (1980), the instrument was modified so that it was appropriate to use with the department heads and within comprehensive institutions. Since the questionnaire had been used, reused, and modified several times by Washburn and its 38

criteria had also been investigated by experts, it was assumed to be valid. With slight modification, it was found appropriate for this research study. A test of reliability was run in the pilot study at Oklahoma State University utilizing the test-retest method. The reliability coefficients were estimated at 1.00 for the rankings of key components and at .81 for the listing of departmental characteristics.

Fifty-three public comprehensive universities and colleges were randomly selected from the list of institutions of higher education prepared by the Carnegie Council on Policy Studies in Higher Education (1976). The selection of subjects for this study was based on Biglan's (1973) three-dimensional model (see Appendix A). One department of each of the eight disciplines was selected to represent that discipline on the basis that it was most likely to exist in randomly selected institutions. The subjects of this study were department heads who were primarily affiliated with the following departments: mathematics, botany, English, psychology, computer science, agricultural economics, economics, and secondary education. Substitutions of the departments within their disciplines were also made.

Four hundred (400) questionnaires were mailed to the subjects. The response rate was 73%, which constituted roughly 10% of the total population. Analysis of data was conducted in four phases, each designed to answer a specific research question. The first phase involved calculating the average percentage of agreement of all respondents in their rankings of five key components (A = community and university service, B = faculty scholarship and research, C = graduate program, D = students, and E = undergraduate teaching). The responses of all respondents who circled the "very important" departmental

characteristics, marked off the "unimportant" ones, and omitted the ones they considered "moderate" were tabulated. Percentages of each characteristic were calculated, and the list in rank-order of the percentages of characteristics was presented (see Table V, Chapter IV). The third phase of data analysis involved the comparison of the top five most important criteria listed in Table V and the criteria identified by department heads of each discipline. The final phase involved calculating the average optimal and minimum acceptable levels of each of the five most important criteria.

Findings

The following findings resulted from the study:

1. Given the five components: community and university service, faculty scholarship and research, graduate program, students, and undergraduate teaching; department heads of public comprehensive universities and colleges ranked faculty scholarship and research as first in importance, graduate program as second and third (see Table IV, Chapter IV), students as fourth, and community and university service as fifth. It is important to note that the undergraduate teaching component was not ranked. Nevertheless, it received a relatively high percentage compared to other components in each category. The analysis, based on key component ranking, replicated the process used in Washburn's (1980) dissertation. If one were to rank key components differently, however (e.g., numerals 5, 4, 3, 2, and 1 could be assigned to replace ranks first, second, third, fourth, and fifth, respectively in Table V), then the sum of numerical values for each key component in all categories would show that department heads

regarded faculty scholarship and research as first in importance, undergraduate teaching as second, graduate program as third, students as fourth, and community and university service as fifth.

2. The single most important criterion that department heads of public comprehensive universities and colleges used in determining departmental excellence was the average number of refereed journal articles published per professor per year.

3. The top five most important criteria for judging departmental excellence were: (a) the average number of refereed journal articles published per professor per year, (b) the average overall student evaluation of departmental courses, (c) the proportion of faculty conducting research or scholarly work with external funding, (d) the proportion of departmental faculty serving on a national review board or professional advisory committee, and (e) the proportion of departmental faculty serving on a university committee.

4. The only similarity of identification of criteria across eight disciplines was that five out of eight disciplines considered the average number of refereed journal articles published per professor per year as the most important criterion in determining departmental excellence.

5. The optimal and the minimum levels of performance of the five most important criteria were as follows:

a. The average optimal level of the number of refereed journal articles published per professor per year was 2.6, while the average minimum level was one refereed publication per professor per year.

b. The average optimal student evaluation score of departmental courses was 6.5; the average minimum score was 2.8.

c. More than 40% of faculty conducting research with external funds was optimal, while 16.5% was the minimum.

d. The proportion of no more than 25% and no less than 10% of departmental faculty should serve on the national review board or professional advisory committee.

e. The proportion of no more than 25% and no less than 10% of departmental faculty should serve on university committees.

6. Although department heads' selections of the most important criterion corresponded with their rankings of its key component (faculty scholarship and research), and more than 6 out of 10 department heads who ranked "faculty scholarship and research" as first also selected the average number of refereed journal articles published per professor per year as the most important criterion, the relationship between the ranking of the key component and the selection of criteria did not occur in other key components.

7. Among the top five most important criteria for judging departmental excellence, four criteria reported in the present study matched those in the Washburn study, which focused on research university departments. These criteria were: (a) the average number of refereed journal articles published per professor per year, (b) the proportion of faculty conducting research or scholarly work with external funding, (c) the proportion of departmental faculty serving on a national review board or professional advisory committee, and (d) the average overall student evaluation of departmental courses.

The average number of refereed journal articles published per professor per year was identified as the most important criterion in both studies. The proportion of faculty conducting research or scholarly work with external funds was listed second in Washburn's study, but was listed as third in the present study; the proportion of departmental faculty serving on a national review board or professional advisory committee was listed as third in Washburn's study, but was listed fourth in the present study. Lastly, the average overall student evaluation of departmental courses was listed last among the top five criteria in Washburn's study, but was listed second in the present study.

The comparison of the results from the two studies indicated the major distinction of emphases in two types of institution: research universities and public comprehensive universities and colleges. While research and publications were emphasized in both types of institution, the emphasis was stronger in research universities than in public comprehensive universities and colleges. Two departmental characteristics found under the faculty scholarship and research component were widely regarded as the most important criteria for judging excellence in research universities. Undergraduate teaching, however, was much less emphasized than faculty publications. Public comprehensive universities and colleges, on the other hand, valued excellence in undergraduate teaching by widely regarding student evaluation of faculty performance in departmental courses as the most important criterion, next to the average number of refereed journal articles published per professor per year.

Conclusions

From the findings of the study, the following conclusions seemed appropriate:

1. For academic departments in public comprehensive universities and colleges to be considered excellent, they must enhance their productivity in the area of faculty scholarship and research, especially with regard to yearly publications and amount of research with external funds. In addition, they must value excellence in undergraduate teaching by emphasizing the student evaluation of faculty performance in departmental courses. The literature review supported this conclusion. Comprehensive universities and colleges have been the synthesizers of the liberal arts orientation and the research orientation. In order to maintain the balance between the two extremes, public comprehensive universities and colleges need to move toward advanced research and instruction. They need to be excellent in both areas.

2. One of the major characteristics of comprehensive universities and colleges was that they were very responsive to the needs and interests of society and the academic community because these two factors were ". . . a counterweight to the academic interests of faculty" (Smith, 1978, p. 473). Therefore, in excellent academic departments of the public comprehensive universities and colleges, there should be a good proportion of departmental faculty serving on national review boards or professional advisory committees, and on university committees.

Recommendations for Further Study

Any researcher attempting to replicate this study should consider the following recommendations:

1. These studies should ideally involve a larger number of participants.
2. These studies might specifically involve smaller public comprehensive universities and colleges that have only a few departments or divisions. Since the responsibilities of these institutions are concentrated within fewer units, the criteria that are used to judge excellence may be different.
3. These studies should modify the instrument used in this study or should develop another instrument that is more precise and consumes a less amount of time in its administration. Many of the respondents felt that the questionnaire was too long. There should be fewer departmental characteristics identified than in the present one. The new instrument should only require respondents to: (a) rank the key components, (b) mark the characteristics they consider very important, and (c) specify the optimal and minimum acceptable levels of activities or achievement on the chosen characteristics.
4. One useful replication of this study would involve criteria used by department heads from private comprehensive universities and colleges in order to compare data with results of this study.
5. Replications of this study should also involve criteria identified by department heads from the liberal arts institutions and the two-year institutions as classified by the Carnegie Council on Policy Studies in Higher Education to compare the emphasis of departments on

excellence in faculty research productivity and on excellence in undergraduate teaching with comprehensive university and research university departments.

Concluding Thoughts

Throughout this study, two issues were of great concern to the researcher. One was the issue of the academic department as an organizational unit, and the other was student evaluation of faculty performance.

Academic Departments

The heritage of departmentalization can be traced back to the medieval period. The University of Salerno and numerous universities established at the end of the Middle Ages in Europe were examples of how masters grouped themselves in the autonomous areas of concentration and set standards for their degrees. In England, the development of separate faculties during the reign of Edward VI was organized at Oxford and Cambridge, but was ceased when both universities came to focus on divinity and liberal arts. In Canada, the University of Toronto sought to maintain the Oxbridge pattern.

The development of departmentalization in England and Canada, according to Andersen (1977), set forth the concept of the department elsewhere. The development of the department was not evident in colonial America until 1739, at Harvard, under the leadership of President Quincy. From here onward the thrust toward American departmentalization was influenced by the development of graduate education,

which made the departments the foundation of the American higher education system.

Modern academic departments were equipped to perform four functions. These were: graduate education, recruitment and promotion of staff members, research, and undergraduate education. With the increasing amount and organization of knowledge, modern academic departments became more specialized than before. Prominent critics in higher education claimed that departmentalization eroded the unity of knowledge. With the greater emphasis on specialization influenced by business and industry, general education lost its ground. Critics noted that academic departments had become very isolated and self-contained. Hutchins (1953) believed that the academic department was an obstacle to institutional and human development.

It was evident that academic departments in comprehensive colleges and universities had exhibited two major weaknesses: lower than expected performance for undergraduate teaching, and low contribution in public service. Due to the emphasis or overemphasis of the department on specialization which involved extensive use of research, teaching at the undergraduate level was ignored. Time and funds were vested largely in research programs and graduate education. Many undergraduate courses were taught by teaching assistants or graduate assistants. Extensive involvement in research made faculty unavailable and inaccessible to the needs of undergraduate students in academic and guidance-counseling areas. Research activities also limited faculty contributions to public services.

Though many institutions attempted to create alternative structures to replace academic department structure, one must realize that

the department has been the central organizational unit of American institutions of higher learning. Over the past century, it has been the vital part of the success of American higher education in providing frontiers of knowledge and educational opportunities. As Trow (1977) noted, it was ironic that when academic reformers in America were attacking the academic department, European universities had been successful in modifying their faculty structures in the direction of discipline-based American departments. Departments, after all, promoted scholarship and protected higher education from outside interference.

Yet, the fundamental weaknesses of the academic department must be remedied. Restoration of a sense of common purpose and defense of general education must take place by pleading to the faculty regarding the importance of training undergraduate students. If undergraduate teaching was not taken seriously, there would be adults who were not well-equipped to do quality graduate study or adults who were not capable of coping with the real world in the job market. Business and industry, in turn, would become less confident in higher education in educating people. A greater number of general education-oriented faculty is needed. Lastly, the public image of research emphasis over teaching emphasis within good universities must also be changed. Teaching must be viewed as respectable as research, and complementary one to the other.

Student Evaluation of Faculty Performance on Departmental Courses

During the past decade, student evaluation has been one of the

most debatable issues in higher education. Research studies had been conducted regarding the subject at all levels of education. Yet, there is still a controversy between what should be done and what is actually done. Student evaluation of departmental courses was ranked as the second most important criterion for determining departmental excellence in this study. Nevertheless, faculty members at comprehensive universities and colleges did not regard the process as personally useful or significant.

To some faculty, students were not mature enough to perform an evaluation that may affect the standing of professor. Student evaluation had been regarded as cursory, unplanned, and capricious. Some suspected that students used evaluation as a tool for punishing or rewarding faculty. As a result, a large proportion of this group of faculty felt that the practice of student evaluations reduced their morale, job satisfaction, and personal confidence in institutional administration.

On the other hand, some faculty believed that students were the only real observers in the classroom. Reports of their classroom experiences provided unique information about the professor and the teaching environment. Instructors could improve their teaching through various approaches, and one way of helping them bring about positive changes in their classroom behavior was to tell them what their students thought of their teaching.

Research studies on the field found empirically both positive and negative effects of student evaluation on faculty attitude and instructional improvement. This might be the reason that the issue remained controversial. Though department heads regarded student evaluation as

very significant, it was questionable to some faculty as to how truthful the evaluation could be. In this researcher's point of view, if student evaluation is to be meaningful, time must be spent in designing valid instruments and in preparing students to evaluate responsibly. These two areas of personal concern to the researcher could result in future studies that might help to shed light on these two controversial areas of concern affecting higher education.

BIBLIOGRAPHY

- Andersen, K. In defense of departments. In: Academic Departments, D. E. McHenry (ed.). San Francisco: Jossey-Bass, 1977, 1-11.
- Alley, L. Two paths to excellence. Quest, 1982, 34, 99-108.
- Astin, A. Equal access to postsecondary education: myth or reality? UCLA Educator, 1977, 19, 8-17.
- Astin, A. and Hensen, J. New measuring of college selectivity. Research in Higher Education, 1977, 6, 1-8.
- Astin, A. and Lee, C. The Invisible Colleges. New York: McGraw-Hill, 1972.
- Astin, A. and Solmon, L. Are reputational ratings needed to measure quality? Change, 1981, 13, 14-19.
- Baird, L. The relationship between ratings of graduate departments. (ERIC Document Reproduction Service, No. ED 203 739.) Princeton, New Jersey: Educational Testing Service, 1980.
- Biglan, A. The characteristics of subject matter in different academic areas. Journal of Applied Psychology, 1973a, 57, 195-203.
- Biglan, A. Relationships between subject matter characteristics and the structure and output of university departments. Journal of Applied Psychology, 1973b, 57, 204-213.
- Brann, J. The chairmen: an impossible job about to become tougher. In: The Academic Department or Division Chairman: A Complex Role, J. Brann and T. Emmet (eds.). Detroit: Balamp, 1972, 5-27.
- Bridgwater, C. and Walkenbach, J. Inaccuracies in recently published departmental rankings: an empirical demonstration. (ERIC Document Reproduction Service No. ED 216 593.) Missoula, Montana: University of Montana, 1981.
- Brown, D. The Mobile Professors. Washington, D.C.: American Council on Education, 1967.
- Brubacher, J. and Rudy, W. Higher Education in Transition. New York: Harper and Row, 1976.

- Carnegie Council on Policy Studies in Higher Education. A Classification of Institutions of Higher Education. Berkeley: Carnegie Foundation for Advancement of Teaching, 1976.
- Carroll, A. The role conflict phenomenon: implications for department chairmen and academic faculty. Improving College and University Teaching, 1976, 24, 245-246.
- Cartter, A. An Assessment of Quality in Graduate Education. Washington, D.C.: American Council on Education, 1966.
- Centra, J. How Universities Evaluate Faculty Performance: A Survey of Department Heads. Princeton: Educational Testing Service, 1977.
- Clark, M., Hartnett, R., and Baird, L. Assessing Dimensions of Quality in Doctoral Education: A Technical Report of a National Study in Three Fields. Princeton, New Jersey: Educational Testing Service, 1976.
- Corson, J. The Governance of Colleges and Universities. New York: McGraw-Hill, 1975.
- Creswell, J., Seagren, A., and Henry, T. Professional development training needs of department chairpersons: a test of a Biglan model. Planning and Changing, 1979, 10, 224-237.
- Dent, P. An analysis of reputational quality ratings in higher education. Educational Research Quarterly, 1978, 4, 21-32.
- DeSouza, A., Vogeler, I., and Foust, B. The overlooked departments of geography. Journal of Geography, 1981, 80, 170-175.
- Dressel, P., Johnson, F., and Marcus, P. The Confidence Crisis: An Analysis of University Departments. San Francisco: Jossey-Bass, 1970.
- Dressel, P. and Reichard, D. The university department: retrospect and prospect. Journal of Higher Education, 1970, 41, 381-402.
- Drew, D. and Karpf, R. Ranking academic departments: empirical findings and theoretical perspective. Research in Higher Education, 1981, 14, 305-320.
- Dubé, W. Undergraduate origins of U.S. medical students. Journal of Medical Education, 1974, 49, 1005-1010.
- Elton, C. and Rodgers, S. The departmental rating game: measure of quantity or quality? Higher Education, 1973, 2, 439-446.
- Endler, N., Ruston, J., and Roediger, H. Productivity and scholarly impact (citations) of British, Canadian, and U.S. departments of psychology. American Psychologist, 1979, 33, 1064-1082.

- Falk, G. The academic department chairmanship and role conflict. ICUT, 1980, 27, 79-86.
- Fretter, W. The Cartter Report on the leading schools of education, law, and business. Change, 1977, 9, 44-48.
- Gardner, J. Excellence. New York: Harper and Row, 1961.
- Gay, L. Educational Research. Columbus, Ohio: Charles Merrill, 1981.
- Gourman, J. The Gourman Report: A Rating of Graduate and Professional Programs in American and International Universities. Los Angeles: National Education Standards, 1977.
- Gourman, J. The Gourman Report: Ratings of American Colleges. Phoenix, Arizona: Continuing Research Institute, 1967.
- Gross, E. and Grambsch, P. University Goals and Academic Power. Washington, D.C.: American Council on Education, 1968.
- Guardo, C. Liberal arts in a comprehensive university. Liberal Education, 1978, 64, 485-495.
- Harclerod, F., Sagen, H., and Molen, C. The Developing State Colleges and Universities: Historical Background, Current Status, and Future Plans. Iowa City, Iowa: American College Testing Program, 1969.
- Harris, J. The pursuit of excellence. (ERIC Document Reproduction Service, No. ED 219 997.) Toledo, Ohio: Toledo University Center for the Study of Higher Education, 1981.
- Hartnett, R., Clark, M., and Baird, L. Reputational ratings of doctoral programs. Science, 1978, 199, 1310-1314.
- Heimler, C. The college departmental chairman. Educational Record, 1967, 158-163.
- Hengstler, D., Bradenburg, D., Braskamp, L., and Amock, H. Faculty ratings as a measure of departmental and administrator quality. Research in Higher Education, 1987, 14, 259-275.
- Hensley, O. How valid are scholars' departmental rating models for assessing the quality of graduate programs? Research in Higher Education, 1980, 12, 255-269.
- Hodgkinson, H. Institutions in Transition. New York: McGraw-Hill, 1971.
- Hughes, R. A Study of Graduate Schools in America. Oxford, Ohio: Miami University Press, 1925.

- Hughes, R. Report of the committee on graduate instruction. Educational Record, 1934, 15, 192-234.
- Hutchins, R. The University of Utopia. Chicago: University of Chicago Press, 1953.
- Jordan, R. Library characteristics of colleges ranking high in academic excellence. College and Research Libraries, 1963, 24, 369-376.
- Keniston, H. Graduate Study and Research in the Arts and Sciences at the University of Pennsylvania. Philadelphia: University of Press, 1959.
- Koelsch, W. Better than thou: the rating of geography departments in the United States, 1924-1980. Journal of Geography, 1981, 80, 164-169.
- Korb, R. Clusters of colleges and universities: an empirically determined system. (ERIC Document Reproduction Service, No. ED 227 797.) Washington, D.C.: National Center for Education Statistics, 1982.
- Krause, E. and Krause, L. The colleges that produce our best scientists: a study of the academic training grounds of a large group of distinguished American scientists. Science Education, 1970, 54, 133-140.
- Kuhn, T. The Structure of Scientific Revolutions. Chicago: University of Chicago Press, 1962.
- Lane, J., Stenlund, H., and Westlund, A. Variety of attitudes towards the comprehensive university. Higher Education, 1982, 11, 441-474.
- Lawrence, J. and Green, K. A Question of Quality: The Higher Education Ratings Game. Washington, D.C.: American Association for Higher Education, 1980.
- Lee, C. Relationship of the department chairman to the academic dean. In: The Academic Department or Division Chairman: A Complex Role, J. Brann and T. Emmet (eds.). Detroit: Balamp, 1972, 54-62.
- Leslie, D. The status of the department chairmanship in the university organization. AAUP Bulletin, 1973, 419-426.
- Levine, A. Handbook on Undergraduate Curriculum. San Francisco: Jossey-Bass, 1978.
- Margulies, R. America's leading professional schools. Change, 1973, 4, 21-27.

- McHenry, D. Academic Departments. San Francisco: Jossey-Bass, 1977.
- Mevey, F. and Hughes, R. Problem of College and University Administration. Cedar Falls, Iowa: Iowa State College Press, 1952.
- Millett, J. The Academic Community: An Essay on Organization. New York: McGraw-Hill, 1962.
- Millett, J. New Structures of Campus Power. San Francisco: Jossey-Bass, 1979.
- Mosier, J. Criteria for excellence of community colleges, as perceived by local community leaders. (Unpub. Doctoral Dissertation, Oklahoma State University, 1983.)
- National Commission on Excellence. A Nation at Risk: The Imperative for Educational Reform. Washington, D.C.: U.S. Department of Education, 1983.
- Roose, K. and Andersen, C. A Rating of Graduate Programs. Washington, D.C.: American Council on Education, 1970.
- Rosenberg, M. The elusiveness of eminence. American Psychologists, 1979, 34, 723-725.
- Rudolph, F. The American College and University: A History. New York: Random House, 1962.
- Smart, J. and Elton, C. Examining departmental management. In: Administrative Roles of Department Chairmen, J. C. Smart and J. R. Montgomery (eds.). San Francisco: Jossey-Bass, 1976.
- Smart, J. and Elton, C. Goal orientations of academic departments: a test of Biglan's model. Journal of Applied Psychology, 1975, 60, 580-588.
- Smith, H. Comprehensive universities and colleges: synthesizers of liberal and professional education. Liberal Education, 1978, 64, 469-484.
- Smith, R. and Fiedler, S. The measurement of scholarly work: a critical review of the literature. Education Record, 1971, 52, 225-232.
- Solmon, L. and Astin, A. Departments without distinguished graduate programs. Change, 1981, 13, 23-28.
- Swartz, J. Errors in citation analysis. American Psychologists, 1979, 34, 723.
- Tidball, M. and Kistiakowski, U. Baccalaureate origins of American scientists and scholars. Science, 1976, 193, 646-652.

- Trow, M. Departments as contexts for teaching and learning. In: Academic Departments, D. E. McHenry (ed.). San Francisco: Jossey-Bass, 1977, 12-33.
- Trow, M. Problems in the transition from elite to mass higher education. Policies for Higher Education. Paris, France: OECD, 1974.
- Tucker, A. Chairing the Academic Department: Leadership Among Peers. Washington, D.C.: American Council on Education, 1981.
- Van Dalen, D. Understanding Educational Research. New York: McGraw-Hill, 1966.
- Waltzer, H. The Job of Academic Department Chairman. Washington, D.C.: American Council on Education, 1975.
- Washburn, B. Standards of departmental excellence: an investigation of the criteria used for judging excellence by selected departmental faculties of a research university. (Unpub. Doctoral Dissertation, Stanford University, 1980.)
- Webster, D. Advantages and disadvantages of methods of assessing quality. Change, 1981a, 13, 20-24.
- Webster, D. How to assess quality in master's programs: a new and better way. (ERIC Document Reproduction Service, No. ED 185 916.) College Park, Maryland: University of Maryland, 1981b.

APPENDIXES

APPENDIX A

BIGLAN'S TABLE

TABLE VIII
CLUSTERING OF ACADEMIC TASK AREAS IN THREE DIMENSIONS

Task Area:	Hard		Soft	
	Nonlife System	Life System	Nonlife System	Life System
Pure	Astronomy	Botany	English	Anthropology
	Chemistry	Entomology	German	Political Science
	Geology	Microbiology	History	Psychology
	Mathematics	Physiology	Philosophy	Sociology
	Physics	Zoology	Russian Communications	
Applied	Ceramic Engr.	Agronomy	Accounting	Educ. Adm. & Super.
	Civil Engr.	Dairy Science	Finance	Second. & Cont. Edu.
	Computer Science	Horticulture	Economics	Spec. Edu.
	Mech. Engr.	Agri. Economics		Voc. & Tech. Edu.

Source: A. Biglan, "Relationships between subject matter characteristics and the structure and output of university departments," Journal of Applied Psychology (1973).

APPENDIX B

QUESTIONNAIRE COVER LETTER

Apiwat Watanangura
704 West 8th, 5E
Stillwater, OK 74074

I am writing to the department heads or chairpersons of the comprehensive universities and colleges for the cooperation in responding to the questionnaire. As a doctoral student at Oklahoma State University, I am conducting a dissertation research concerning the criteria that department heads use in determining departmental excellence. Research study that concerns the identification of criteria for determining departmental excellence at the comprehensive universities and colleges has been lacking. Department heads--as leaders who set the pace and maintain standards of the departments--are, therefore, very important participants in this research study, as they identify the criteria they use for determining excellence of the departments in their disciplines.

This dissertation research is being conducted under the direction of Dr. John J. Gardiner, Associate Professor of Educational Administration and Higher Education. Dr. Robert B. Kamm, Dr. William E. Segall, and Dr. Lynn K. Arney are also members of the committee. Dr. Bill H. Washburn, who developed the original questionnaire at Stanford University, has permitted me to modify and replicate his questionnaire. He has also been very supportive toward this dissertation research.

With your response, I expect to develop a classification of criteria that will be useful for determining departmental excellence. Then, the comparison between and among disciplines will be made to examine the similarities and/or differences of criteria identified by department heads. Your responses will be treated with strict confidentiality; neither departments nor department heads will be identified in reporting the results.

Attached is a copy of the questionnaire which should take no more than 30 minutes to complete. I hope you will respond to the questionnaire and mail it back by May 15, 1984. If a delay cannot be avoided, please return it as soon after May 15 as possible. Please use the enclosed stamped, self-addressed envelope for returning the questionnaire.

Your kind consideration and cooperation are immensely appreciated.

Respectfully,

Apiwat Watanangura

APPENDIX C

QUESTIONNAIRE

**CRITERIA IN DETERMINING
DEPARTMENTAL EXCELLENCE**

Modified from the Study of Criteria
and Standards of Departmental
Excellence: Faculty Questionnaire
Invented and Developed by
Bill H. Washburn at Stanford University

DIRECTIONS

This questionnaire consists of a table and a few short questions. *Table of Departmental Excellence* (pp. 3-5) is divided into five categories (designated by letters A to E). Within each substantive category, relevant departmental characteristics are listed and identified by number.

Completing the Table entails three tasks:

- (1) Ranking the categories of departmental activity from most to least important;
- (2) Circling the most important and crossing-out the least important characteristics;
- (3) Specifying optimal and minimum acceptable levels of activity or achievement.

FIRST STEP Review the five *categories* of departmental characteristics shown on pages 3-5 and then:

- *rank the categories* (1 = most important; 5 = least important) according to their importance for evaluating departmental excellence. Record the rank in the adjoining box. When categories are equally important, give them the same rank.

SECOND STEP Read over the departmental *characteristics* listed within each category and then:

- *cross-out the characteristics* which you consider unimportant in the evaluation of departmental excellence in your discipline.
- *circle the number of each characteristic* which you consider very important for judging departmental excellence in the discipline.
- *write-in omitted characteristics* (in the appropriate category) which you think are very important for assessing departmental excellence.

THIRD STEP Considering *only* those characteristics you circled, start with the first characteristic in the category now ranked '1' and:

- *record the number or percentage** which you think reflects an optimal level of activity or achievement. (Use the column labeled 'optimal' as shown in the illustration.)
- *record the number or percentage** which you think represents a minimum acceptable level of activity for distinguished departments in the discipline. (Use the column labeled 'minimum acceptable' as shown in the illustration.)
- *specify the optimal and minimum-acceptable** levels of activity or achievement for each of the remaining circled characteristics on the Table and record the figures in the appropriate places.
- *specify the optimal and minimum acceptable activity levels* for the items you added to the Table.*

* If assigning a number or percent seems inappropriate for any particular characteristic, just write 'NA' instead. If you have no informed opinion about optimal or minimum acceptable level of activity for a given characteristic, put a '?' in place of a number.

- - - SEE THE INSERT FOR AN ILLUSTRATION OF THE STEPS - - -

ILLUSTRATION

- a. If you consider category A (COMMUNITY AND UNIVERSITY SERVICE) the single most important area for determining departmental excellence, place a '2' in the adjacent box.
- b. If you consider the first characteristic in category A very important in judging departmental excellence, circle the numeral '1' which designates that characteristic.
- c. If you think the second characteristic is unimportant for evaluating departmental excellence, cross it out.
- d. If you think the third characteristic under category A is neither unimportant nor very important in determining excellence, then pass it by and proceed to the next characteristic.
- e. If you think the fourth characteristic is very important, circle the '4' next to it.
- f. If you have no particular opinion about the relative importance of the fifth listed characteristic, leave it alone.
- Two of the characteristics (1 and 4) require further attention.
- g. On characteristic '1,' if 33% is the optimal level of activity and 10% is the minimum acceptable level of activity, those figures should be written in the appropriate columns.
- h. If, on characteristic '4,' you think an optimal level of activity cannot be appropriately specified, write 'NA' in the left column. If you have no informed opinion about the minimum acceptable level of activity on the characteristic, place a '?' in the right column.

rank			MINIMUM	
			OPTIMAL	(scale) ACCEPTABLE
2	A	COMMUNITY AND UNIVERSITY SERVICE		
1		proportion of departmental faculty serving on a university committee.....	33%	0 - 100% 10%
2		proportion of departmental faculty elected to the faculty senate.....		0 - 100%
3		proportion of departmental faculty serving on a national review board or professional advisory committee.....		0 - 100%
4		average number of consulting days per professor per quarter.....	NA	number ?
5		proportion of faculty who serve as general advisors for undergraduates without a major.....		0 - 100%
6				
7				

TABLE OF DEPARTMENTAL EXCELLENCE

rank								
	<input type="checkbox"/> A	COMMUNITY AND UNIVERSITY SERVICE	OPTIMAL	(scale)	MINIMUM	ACCEPTABLE		
1		proportion of departmental faculty serving on a university committee.....		0 - 100%				
2		proportion of departmental faculty elected to the faculty senate.....		0 - 100%				
3		proportion of departmental faculty serving on a national review board or professional advisory committee.....		0 - 100%				
4		average number of consulting days per professor per quarter.....		number				
5		proportion of faculty who serve as general advisors for undergraduates without a major.....		0 - 100%				
6								
7								
	<input type="checkbox"/> B	FACULTY SCHOLARSHIP AND RESEARCH	OPTIMAL	(scale)	MINIMUM	ACCEPTABLE		
8		proportion of faculty who have received a "top prize" in the discipline.....		0 - 100%				
9		proportion of faculty who have received an honorific or competitive fellowship.....		0 - 100%				
10		proportion of faculty elected to a national or professional honorary society.....		0 - 100%				
11		average number of refereed journal articles published per professor per year.....		number				
12		average number of books (non-text) published per professor per two years.....		number				
13		average number of textbooks published per professor per two years.....		number				
14		proportion of faculty serving as a journal editor or referee.....		0 - 100%				
15		proportion of faculty conducting research or scholarly work with external funding.....		0 - 100%				
16		average number of citations per professor in the latest citation index.....		number				

B FACULTY SCHOLARSHIP AND RESEARCH (cont.)		OPTIMAL	(scale)	MINIMUM ACCEPTABLE
17	proportion of faculty with rank of "full" professor		0 - 100%	
18	proportion of faculty who are involved in interdepartmental programs		0 - 100%	
19				
20				

rank <input type="checkbox"/> C GRADUATE PROGRAM		OPTIMAL	(scale)	MINIMUM ACCEPTABLE
21	proportion of admitted applicants who actually enroll in the graduate program		0 - 100%	
22	proportion of entering graduate students who drop-out by the end of the first year		0 - 100%	
23	average number of graduate students per professor		number	
24	average number of master's degrees awarded per professor per year		number	
25	proportion of master's degrees students who complete the program		0 - 100%	
26	proportion of the students in the department who are graduate students (balance between undergrad and graduate students)		0 - 100%	
27	proportion of graduate students who publish their work while in the program		0 - 100%	
28	average number of master's degrees awarded in the department each year		number	
29				
30				

rank					
	<input type="checkbox"/> D	STUDENTS	OPTIMAL	(scale)	MINIMUM ACCEPTABLE
31		average SAT Verbal score of undergraduate majors		200 to 800	
32		average SAT Math score of undergraduate majors		200 to 800	
33		proportion of undergraduate majors elected to Phi Beta Kappa		0 - 100%	
34		proportion of undergraduate majors who graduate and pursue graduate training in the discipline		0 - 100%	
35		average GRE Verbal score of entering graduate students		200 to 800	
36		average GRE Math score of entering graduate students		200 to 800	
37		proportion of graduate students who have external fellowship funding		0 - 100%	
38		proportion of graduate students enrolled as full-time students		0 - 100%	
39					
40					

rank					
	<input type="checkbox"/> E	UNDERGRADUATE TEACHING	OPTIMAL	(scale)	MINIMUM ACCEPTABLE
41		average number of undergraduate majors per professor		number	
42		proportion of departmental faculty who teach at least one undergraduate course		0 - 100%	
43		proportion of faculty who advise undergraduate majors in the department		0 - 100%	
44		average overall student evaluation of departmental courses		1 - 7	
45		proportion of faculty who have won a prize for teaching in the last 5 years		0 - 100%	
46		proportion of undergraduate courses developed or redesigned in the last 3 years		0 - 100%	
47					
48					

49. If you wish, please use this space to express your views regarding the level of departmental excellence and/or the nature of academic standards.

50. What accomplishments tend to bring the department national recognition of excellence in its discipline?

51. Which departmental characteristics seem to be, as best you can tell, the ones University administrators count most heavily in judging departmental strength?

52. With which department in the University are you primarily affiliated?

53. Nationally, what is the department's present standing in its discipline?

Number One	_____
Top Three	_____
Top Five	_____
Top Ten	_____
Top Fifteen	_____
Top Twenty	_____
Not sure	_____
Other _____	_____

(specify)

APPENDIX D

FOLLOW-UP LETTER TO NONRESPONDENTS

Apiwat Atanangura
704 West 8th, 5E
Stillwater, OK 74074

Three weeks ago I sent you a letter requesting your cooperation in responding to the Criteria in Determining Departmental Excellence Questionnaire. To date, I have not received your response.

I realize that you have many responsibilities to fulfill. However, I would like to encourage you to complete the questionnaire and return it to me in the enclosed, self-addressed, postage-paid envelope as soon as possible. The topic of my dissertation is focused on the criteria that department heads use in determining departmental excellence. Your kind consideration and cooperation are imperative to the successful completion of this study.

If you have any questions regarding the questionnaire or your participation, I can be reached at (405) 624-1303.

Thank you for your kindness.

Respectfully,

Apiwat Watanangura

APPENDIX E

TABLES SHOWING RANKING OF KEY COMPONENTS

TABLE IX
RANKING OF KEY COMPONENTS IN PURE-HARD
NONLIFE SYSTEM DISCIPLINE*

Component/Rank	First (%)	Second (%)	Third (%)	Fourth (%)	Fifth (%)
Community and University Service	--	--	--	29	71
Faculty Scholarship and Research	43	28	29	--	--
Graduate Program	--	14	29	43	14
Students	10	29	32	29	--
Undergraduate Teaching	39	33	28	--	--

*Represented by Astronomy, Chemistry, Geology, Mathematics, and Physics

TABLE X
RANKING OF KEY COMPONENTS IN PURE-HARD
LIFE SYSTEM DISCIPLINE*

Component/Rank	First (%)	Second (%)	Third (%)	Fourth (%)	Fifth (%)
Community and University Service	--	--	30	5	65
Faculty Scholarship and Research	75	--	5	20	--
Graduate Program	--	25	50	--	25
Students	25	65	--	10	--
Undergraduate Teaching	15	25	--	50	10

*Represented by Botany, Entomology, Microbiology, Physiology, and Zoology

TABLE XI
RANKING OF KEY COMPONENTS IN PURE-SOFT
NONLIFE SYSTEM DISCIPLINE*

Component/Rank	First (%)	Second (%)	Third (%)	Fourth (%)	Fifth (%)
Community and University Service	--	--	20	40	40
Faculty Scholarship and Research	60	--	20	20	--
Graduate Program	--	20	80	--	--
Students	--	40	10	50	--
Undergraduate Teaching	20	60	20	--	--

*Represented by English, German, History, Philosophy, Russian, and Communications

TABLE XII
RANKING OF KEY COMPONENTS IN PURE-SOFT
LIFE SYSTEM DISCIPLINE*

Component/Rank	First (%)	Second (%)	Third (%)	Fourth (%)	Fifth (%)
Community and University Service	--	--	27	18	55
Faculty Scholarship and Research	55	36	--	--	9
Graduate Program	--	37	27	18	18
Students	9	9	36	36	10
Undergraduate Teaching	36	27	9	28	--

*Represented by Anthropology, Political Science, Psychology, and Sociology

TABLE XIII
RANKING OF KEY COMPONENTS IN APPLIED-
HARD NONLIFE SYSTEM DISCIPLINE*

Component/Rank	First (%)	Second (%)	Third (%)	Fourth (%)	Fifth (%)
Community and University Service	--	--	20	20	20
Faculty Scholarship and Research	60	20	--	--	--
Graduate Program	--	60	--	20	--
Students	--	40	60	--	--
Undergraduate Teaching	40	--	20	30	--

*Represented by Ceramic Engineering, Civil Engineering, Computer Science, and Mechanical Engineering

TABLE XIV
RANKING OF KEY COMPONENTS IN APPLIED-
HARD LIFE SYSTEM DISCIPLINE*

Component/Rank	First (%)	Second (%)	Third (%)	Fourth (%)	Fifth (%)
Community and University Service	--	15	25	25	35
Faculty Scholarship and Research	50	--	50	--	--
Graduate Program	25	75	--	--	--
Students	--	25	--	50	25
Undergraduate Teaching	75	--	25	--	--

*Represented by Agronomy, Dairy Science, Horticulture, and Agricultural Economics

TABLE XV
RANKING OF KEY COMPONENTS IN APPLIED-
SOFT NONLIFE SYSTEM DISCIPLINE*

Component/Rank	First (%)	Second (%)	Third (%)	Fourth (%)	Fifth (%)
Community and University Service	--	20	--	20	60
Faculty Scholarship and Research	80	20	--	--	--
Graduate Program	--	80	20	--	--
Students	--	20	20	40	20
Undergraduate Teaching	20	--	50	20	--

*Represented by Accounting, Finance, and Economics

TABLE XVI
RANKING OF KEY COMPONENTS IN APPLIED-
SOFT LIFE SYSTEM DISCIPLINE*

Component/Rank	First (%)	Second (%)	Third (%)	Fourth (%)	Fifth (%)
Community and University Service	--	37	12	12	37
Faculty Scholarship and Research	24	--	37	49	--
Graduate Program	12	24	49	12	--
Students	24	12	12	37	24
Undergraduate Teaching	61	49	--	--	--

*Represented by Educational Administration and Supervision, Secondary and Continuing Education, Special Education, and Vocational and Technical Education

2
VITA

Apiwat Watanangura

Candidate for the Degree of

Doctor of Education

Thesis: DEPARTMENTAL EXCELLENCE: CRITERIA USED BY DEPARTMENT
HEADS AT PUBLIC COMPREHENSIVE UNIVERSITIES AND COLLEGES

Major Field: Higher Education

Biographical:

Personal Data: Born in Bangkok, Thailand, June 13, 1957, the
son of Archeep and Poonsri Watanangura.

Education: Graduated from Bovornives High School, Bangkok, Thailand, in May, 1974; received Bachelor of Education degree from Chulalongkorn University with a major in Teaching English as a Foreign Language, in March, 1979; received Master of Education degree from Loyola College, with a major in Educational Management and Supervision in May, 1981; completed requirements for the Doctor of Education degree at Oklahoma State University in December, 1984.

Professional Experience: Project Officer in the Evaluation Division of the Curriculum Development Center, Ministry of Education, Bangkok, Thailand, 1979-80.