8504747

Ashton, Arthur Benner

A RESEARCH ACTIVITY INDEX OF MAJOR RESEARCH UNIVERSITIES

The University of Arizona PH.D. 1984

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

A RESEARCH ACTIVITY INDEX OF MAJOR RESEARCH UNIVERSITIES

by

Arthur Benner Ashton

A Dissertation Submitted to the Faculty of the

CENTER FOR THE STUDY OF HIGHER EDUCATION

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

1 9 8 4

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Final Examination Committee, we certify that we have read the dissertation prepared by Arthur B. Ashton

entitled A Research Activity Index of Major Research Universities

and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

Larry L Leslie — Date 9/27/84
Larry L. Leslie

Arthur T. Grant — Date 9/27/84
Arthur T. Grant

Don L Bowen — Date 9/27/84
Don L. Bowen

Date

Date

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copy of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

Larry L Leslie — Date 9/27/84
Dissertation Director

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and deposited in the University Library to be made available to borrowers under the rules of the library.

Brief quotations from this dissertation are allowed without special permission, provided that acknowledgment of the source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: Arthur B. Ashton

ACKNOWLEDGMENTS

I thank the teachers and mentors that have encouraged me to pursue education. I can only hope that my work will justify their confidence. I particularly acknowledge Professor Larry L. Leslie, chairman of my committee and Professors Don L. Bowen and Arthur T. Grant the other members of my committee for their support and encouragement. In addition, Dr. James J. Rusk was of assistance in my early efforts to identify my dissertation topic.

With every gain there is sacrifice; the magnitude of both are difficult to determine. I extend my thanks to my wife Joanne and my daughters Kacy and Donna for their sacrifices and I can hope that they share in any gains.

TABLE OF CONTENTS

LIST OF TABLES vi

ABSTRACT vii

1. INTRODUCTION 1

Organization of Study 2
Problem Statement 2
Background 3
Purpose 8
Significance 9
Definition of Terms 10
Assumptions 12
Scope and Limitations 12
Summary 14

2. REVIEW OF LITERATURE 15

Indices 15
Related Studies 23
Summary 30

3. DATA SOURCES AND RESEARCH DESIGN 32

Research Questions 32
Analytical Framework 32
Data Sources 35
    Population 35
    Criteria 38
    Primary Data Sources 39
    Variables 41
Method 46
Special Limitations 56
Summary 56

TABLE OF CONTENTS--Continued

4. FINDINGS 58

Development of RAI 58
Data Reduction 58
Selection of Principal Components Method .. 64
RAI Scores 67
Validity 74
Interpretation of Components 74
Analysis of Variables impacts on Rankings . 77
Analysis of Institutional Rankings 84
Consultants Review 88
Reliability 89
Summary 91

5. SUMMARY AND IMPLICATIONS 93

Summary of Framework 93
Summary of Research Design 94
Summary of Findings 96
Development of RAI 96
Validity of RAI 97
Reliability of RAI 101
Implications 101

APPENDIX A: POPULATION OF UNIVERSITIES 109

APPENDIX B: QUALIFICATIONS OF DATA 112

APPENDIX C: CORRELATION COEFFICIENTS........... 116

REFERENCES 118

LIST OF TABLES

Table		Page
2.1	Changes in Dependency on Federal Funds for Support of R&D	25
3.1	Population of Universities-Ranked According to R&D Expenditures	36
3.2	Research Variables	43
4.1	Final Research Variables	63
4.2	Comparison of the 1st PC Ranking to the Rankings of Combined Components	66
4.3	Research Activity Index Ranking-First Principal Component Method	68
4.4	Research Activity Index Ranking-Rotated First Component Method	70
4.5	Comparison of R&D Expenditures, RAI Rotated, and RAI First Principal Component Rankings	72
4.6	Principal Component Analysis Score Coefficients	75
4.7	Rotated Component Analysis Score Coefficients	76
4.8	Selected Institutions for Analysis of Variable Values	79
4.9	Selected Institutions With Large RAI Ranking Differences	86
4.10	Comparison of RAI Based Rankings Without and With the E&G Expenditure Variable	91

ABSTRACT

The purpose of this study was to determine if a composite Research Activity Index (RAI) score could be developed for each major research university in the United States. Such an index if more valid and reliable than existing univariate rankings, would be a valuable instrument for the objective measurement and dissemination of research activity information.

Composite RAI scores were developed for each major research university included in this study. The primary data sources were the National Science Foundation's University Science Statistics and the Association of Research Library's ARL index. A wide range of variables were reduced to eleven variables that constituted the ultimate RAI measure. These variables were as follows: average percentage change in research and development expenditures over four years; total research and development expenditures; total capital expenditures for scientific and engineering facilities and research equipment for research, development, and instruction; full-time scientists and engineers employed; part-time scientists and engineers employed; full-time graduate science students enrolled; part-time graduate science students enrolled; postdoctoral

personnel; other non-faculty doctoral research staff/students; Phd's awarded per year; and the Association of Research Library's annual index. Principal component analysis, a branch of factor analysis, was used as a descriptive statistic to transform linearly the final set of variables into a new set of variables called the first principal component. The first principal component variable loadings were standardized then multiplied by the corresponding standardized data values for each research university. These were combined linearly to produce the RAI scores.

The RAI was validated by examination for systematic bias, by analysis of each variable's contribution to the outcome, by comparison to the traditional research and development expenditures ranking, and by having the statistical methods reviewed by a noted statistician. Reliability was established by reviewing the stability of the data bases and variable's definitions over time, by reviewing their stability in previous studies, and by determining that like universities produced like RAI scores.

This study demonstrates that an objective composite RAI can be developed from existing data and that the index is more valid and reliable than current, unitary research measures. The RAI has the potential for assisting national policy analyses and university management, strategic planning, and evaluation. The RAI allows for historical, longitudional, and trend analysis; comparative analysis on a national, regional, state, or university basis; and the potential setting of objective, future research goals with subsequent evaluations based upon actual performance of RAI scores.

CHAPTER 1

INTRODUCTION

Research and development is directly related to the advancement of society. It is an instrument of federal, and to a lesser extent state policy. Research is central to university aspirations of excellence, prestige, and influence. It is a primary mission of universities and supports the other missions of (graduate) student instruction and public service.

Governmental public policy and university management decision making require comprehensive, valid, and reliable indicators of research activity. This study proposes to develop a new indicator titled, the Research Activity Index (RAI), a measure of research activity in the major research universities. The primary research objectives are to

1. develop criteria for the selection of variables that are measures of research activity,
2. develop a composite Research Activity Index score for each research university through the application of statistical procedures,
3. rank order the universities, using the composite RAI, and
4. evaluate the strengths and weaknesses of the RAI in terms of its validity and reliability as compared to traditional measures of research.

Organization of the Study

The first chapter identifies the societal interest in the current status of, and proposed solutions to, the need for research indicators. The concept of a Research Activity Index is presented along with its potential benefits. In Chapter Two the literature related to research indicators in higher education is reviewed. Chapter Three presents the research design and the data sources. In chapter four the findings are presented. The final chapter summarizes the study, and presents conclusions which examine the implications of the Research Activity Index.

Problem Statement

The development of indicators that can be used for policy analysis and management decision making is a major concern of Congress and educators (NSB 1978, p. viii), particularly the need to develop composite, as well as, component indicators (McCoy, Jones and Makowski 1979, p. 65). Composite indicators are for an entire university; component indicators are for a sector of a university such as each field of science.

Currently, the indicators used to examine research in higher education are univariate; they are single variables such as research and development expenditures, number of graduate level personnel or number of graduate students. These are useful indicators, but it is frequently up to the interpreter of such indicators to analyze and integrate these bits and pieces of data to arrive at a comprehensive understanding of research activity. This study proposes to develop a composite index of research activity for each university. This index could provide decision makers with a more comprehensive indicator for policy analysis and management decision making.

Background

Indicators of higher education are used by federal, regional, state, and institutional level organizations. The more important uses and users are reviewed in this background section.

Since 1972, the National Science Foundation (NSF) has been developing a series of science indicators in an attempt to assess the state of science. Numerous input-related indicators have been developed, as well as a number of creative output-related indicators. For example, citation indices, patents granted, royalities and fees, and technological innovation indicators have been created. In its Science Indicators publication, the National Science Foundation suggested that an adequate conceptual model for research and development is needed. The Foundation notes that the development and tracking of a wide range of indicators over time might help in developing such a model

(National Science Board 1976, p. viii).

The bases for many of the science indicators are the annual National Science Foundation surveys of higher education institutions (National Science Foundation 1979). These surveys provide data on sources of R&D funds and expenditures, federal support, personnel, graduate students and postdoctoral fellows. The importance and uses of these data are noted in a 1979 report to the President and Congress:

> Universities and colleges in the United States, the main source of the nation's scientists and engineers, perform over one-half of the nation's basic research and nearly one-half of Federally sponsored basic research. The level, distribution, and nature of Federal academic obligations are, therefore, of interest to officials at the state and local levels and in nongovernmental sectors as well as to those in the Federal government. Academic officials rely on data in this survey for information on an institution's progress over time and for comparisons with other institutions (National Science Foundation 1979, p. v.).

National Science Foundation survey data are generally presented in annual and time-series tabulations. These are straightforward presentations with little data reduction work done. For example, total research and development expenditures at higher education institutions are rank-ordered from the largest to the smallest dollar amounts for the top 100 research universities. The users of such data must factor in additional information if they want to arrive at a comprehensive view of research activity in universities.

The National Center for Educational Statistics (NCES) like the National Science Foundation, has a Congressional mandate to monitor and report on the condition of education. Through the National Center for Higher Education Management Systems (NCHEMS), the National Center for Educational Statistics has supported a major project to develop higher education indicators. These indicators, much as the Consumer Price Index (CPI) monitors consumer price changes, are to monitor the status and progress toward accomplishing such goals as student choice, institutional productivity, student access, and research (McCoy, et al. 1979, pp. 63-75).

Regional accreditation associations use indicators to evaluate each institution that is a candidate for accreditation or reaffirmation. William T. Haywood reports that the Commission on Colleges of the Southern Association of Colleges and Schools requires schools to conduct a comprehensive self-study using standards and criteria established by the Commission (1980, p. 17).

Financial health indicators are used at the state-level for annual operational and long-range strategic decision making. The important indicators are trends, relationships among trends, and comparisons (Schmidtlein and Lapovski 1980, p. 19). In comparison to the federal government, states do not fund research extensively; they have a lesser, though significant interest in research activity, particularly in regard to agricultural research in the land grant universities.

Research indicators have a variety of uses at the university level. For example, the University of Arizona publishes a report on its annual research and development expenditure rankings which are compared to peer, as well as, to all research universities. The report also indicates the rankings within fields of science. In addition to rankings, tables are prepared that compare universities on their percentage of total research dollars expended. These analyses are used to establish broad goals and monitor progress towards goals (Research and Development at Major Universities Between 1975 and 1979, pp. i-ii, 1-75).

Considerable attention in recent decades has been given to the development of quality indicators of research doctorate programs. One of the latest of these studies is titled, An Assessment of Research--Doctorate Programs in the United States: Mathematics and Physical Sciences, (Jones, Lindzey and Coggeshall, eds., 1982) sponsored by the Conference Board of the Associated Research Council. This report notes that "It is evident that the assessment of graduate programs is highly important for university administrators and faculty, for employers in industrial and government laboratories, for graduate students and prospective graduate students, for policymakers in state and national organizations, and for private and public funding agencies." (Jones, et al., 1982, pp. 1-2). The RAI indicator developed in this study does not measure program quality but undoubtedly has a qualitative dimension. It could be used in ranking research universities and university fields of science without the controversy that surrounds the judgmental rankings.

The interest in indicators, particularly a RAI, may be in direct relationship to the financial commitment involved. The greatest interest appears to be on the part of the federal government and individual universities. National organizations, such as the American Council on Education, the National Association of College and University Business Officers, and the National Center for Higher Education Management Systems should also have an interest in a RAI because of the constituencies they serve.

Presently, the best known and most widely used index of research activity at universities is the simple ranking by dollar amount of research and development expenditures. This unidimensional ranking does not take into account a number of readily available variables that may also help explain research activity. For example, such variables as the number of graduate faculty, graduate students, and average faculty salaries can add multiple dimensions to help explain research activity.

A Research Activity Index for each field of science (RAI-FS), may be also possible. In practice, a RAI-FS may be more useful than a RAI because the former breaks down the RAI into manageable, analytical components of fields of science. For example, employment decisions by faculty and attendance decisions by graduate students are more likely to be related to research activity in a given field of science than in the standing of an entire university. The RAI-FS will not be developed in this study, but if the university level RAI proves beneficial, then its concept and method can be applied to a RAI-FS.

Purpose

How can research activity in universities be measured in a more comprehensive manner than the present simple counting and ranking of research and development dollars expended? The purpose of this study is to develop a composite measure that is a valid and reliable means of measuring research activity at the major research universities.

Significance

This study is significant because of the large number of potential beneficiaries and the importance of the activities in which the RAI might be used for analytical purposes. The following paragraphs outline who may benefit, how benefits may be realized, and the forms of analyses for which a RAI could be used.

The major beneficiaries of a RAI could be (1) federal organizations such as the National Science Foundation and the National Center for Educational Statistics; (2) higher education organizations like the National Center for Higher Education Management Systems, National Association of College and University Business Officers (NACUBO), Education Commission of the States (ECS), American Council on Education (ACE), and accreditation associations; (3) research universities; and (4) prospective faculty members and students.

The beneficiaries could use the RAI to help support (1) national policy analysis, (2) university planning and management, (3) resource acquisition and allocation, (4) monitoring research status and progress towards goals, and (5) classification of research universities.

The forms of analyses could include changes and trends in rankings. Changes in RAI scores could be examined to determine changes in levels of research activity (quite aside from changes in rank orders).

Definition of Terms

Composite Indicator--a global measure based upon several variables.

Fields of Science--

- Engineering
  - Aeronautical & Astronautical
  - Chemical
  - Civil
  - Electrical
  - Mechanical
  - Metallurgy and Materials
  - Other
- Environmental Science
  - Atmospheric Sciences
  - Geological Sciences
  - Oceanography
  - Other
- Life Sciences
  - Biological Sciences
  - Agricultural Sciences
  - Medical Sciences
  - Other
- Mathematical and Computer Sciences

Mathematics

Computer Sciences

Other

Physical Sciences

Astronomy

Chemistry

Physics

Other

Psychology

Social Sciences

Anthropology

Economics

History

Linguistics

Political Science

Sociology

Other

Other Sciences

Research and development expenditures--all expenditures in a fiscal year for basic, applied, and developmental research.

Research Activity Index Score--the sum of the products of the standardized principal component values for each variable and the value of each standardized variable for a given university.

Assumptions

Research, particularly basic research in higher education institutions in the United States, is largely a federally supported activity designed to support national policies. For this reason the dominant use for the RAI is federal policy analysis. An important secondary use is to support planning and management processes within research universities. Both uses assume rationally derived decision making based upon accurate and objective data.

Scope and Limitations

The scope of this study is limited to research performed in the top 100 research universities within the United States. The data that are used in this study have been collected for a number of years, but only 1980 data will be analyzed. RAIs for each field of science will not be developed in this study. Time and cost are the major considerations for setting this scope.

From a federal policy level perspective, an obvious limitation is that the RAI encompasses only research activities in research universities within the United States. Research activities of other research institutions, government agencies, and the private sector are not included, nor are research activities in foreign countries.

This study is also limited in that National Science Foundation surveys, which are the primary data sources for this study, do not collect research and expenditure data for the fields of education, law, humanities, music, arts, physical education, library sciences, and all other non-science fields. This exclusion limits the comprehensiveness of the study. It may be possible to use the Higher Education Generalized Information Survey (HEGIS) research data for these non-science fields. Since research in these fields is relatively small, it is doubtful that the RAI would change significantly if these fields were included in this study. The desirability and feasibility of including non-science data will be left to future studies.

Research and development expenditures are defined by the National Science Foundation as separately budgeted research and development. A few universities separately budget their departmental research while most do not. To the extent that research and development expenditures are separately budgeted in some universities, their RAI score will be overestimated. Of course, this is also true of the research and development expenditure based rankings. There is no reasonable way to correct for these differences in reporting of research and development expenditures.

Another limitation is that the variables selected for this study are limited to those readily available. There are other relevant variables, but time, financial considerations and serious questions as to their validity and reliability have ruled them out. If the results of this study appear promising, most of the above limitations could be overcome in future studies.

Summary

Most measures of research activity are univariant indicators, such as, research and development expenditures or numbers of research faculty. These tend to be inadequate measures of overall research activity. This study proposes to develop a composite indicator of research activity that will provide a more valid and reliable measure of research activity than any univarient measure currently in use.

CHAPTER 2

REVIEW OF LITERATURE

This chapter examines previous research involving indices related to higher education and recent research using methods and variables that are related to the development of the RAI. Literature related to the history of, and the need for, indicators is covered in Chapter One.

Indices

In recent years numerous rankings of higher education institutions and their programs have been performed. A few have drawn attention, primarily because they were sponsored by nationally prominent organizations. The American Council on Education (ACE) sponsored An Assessment of Quality in Higher Education by Allan Cartter (1966); then a more refined, subsequent study titled A Rating of Graduate Programs by Kenneth D. Roose and Charles J. Anderson (1970) was performed. Most recently An Assessment of Research Doctorate Programs in the United States: Mathematics and Physical Sciences, edited by Lyle V. Jones et al., (1982) was one of five such studies sponsored by The Conference Board of Associated Research Councils. The first two of these studies have been criticized because they are thought to measure quality of academic programs using subjective indicators (Gerhan 1979, p. 15; Dolan 1976, p. 3).

These subjective indicators were opinions derived from reputational surveys of faculty, deans and others in higher education. However, different rankings have been performed which are objective and descriptive rather than subjective and prescriptive (Jones et al., 1982, p. 1; Mandel & Johnson 1981, p. 5). This paper will differentiate between the subjective and objective rankings with the goal of demonstrating that the RAI can be an objective index not to be confused with the subjective indices.

Cartter's study covered doctoral programs at 106 institutions involving 29 fields of study. Some 5,367 academic professionals were asked to rank institutions by disciplines according to the quality of graduate faculty and the effectiveness of their doctoral programs. Quality of faculty was graded as distinguished, strong, good, adequate and not graded. Institutions were ranked within the distinguished and strong categories. Reactions to the Cartter and the subsequent and similar Roose-Anderson study were at times bitter although the studies have become widely used reference works. One of the more outspoken opponents of such rankings has been William A. Arrowsmith, who in his preface to W. Patrick Dolan's book, The Ranking Game: The Power of the Academic Elite, says,

> They [the ACE Reports] purport to represent "expert opinion"--that is, the judgment of the professoriate on its own performance. The rankings doubtless indicate something, but as an index of quality they are of extremely dubious value; in fact, they are little more than quantified gossip or hearsay (1976, p. viii).

Dolan's book is devoted to questioning the assumptions, methods, and contents of these two studies. Additional criticisms have come from Gerhan (1979), Glower (1980), and Webster (1981).

In spite of these criticisms, a third major study has been produced because a number of organizations believed the benefits of such work more than outweighed the negative effects. The new study was sponsored by the Conference Board of Associated Research Councils. The first of five publications was titled, An Assessment of Research-Doctorate Programs in the United States: Mathematical and Physical Sciences (1982); four other reports covered other fields of science. The Conference Board consists of the American Council on Education, the National Research Council, and the Social Science Research Council. Financial support was provided by the Andrew W. Mellon Foundation, the Ford Foundation, the Alford P. Sloan Foundation, the National Institute of Health, the National Science Foundation, and the National Academy of Sciences

(1982, p. iii). The study was built upon criticisms of earlier studies. It encompassed 2,699 graduate programs in thirty-two diciplines, both subjective and objective variables were used. A multidimensional approach to the assessment of doctorale programs was done to avoid the problems associated with using only judgmental ratings. Data for sixteen variables were arrayed but were not combined in any manner. Also, in the development of peer rating variables evaluators were provided with the names of faculty members in each program to be rated, and the number of research doctorates awarded in the last five years. These approaches vary from those in the previous studies (Jones et al., eds., 1982, p. 1).

David S. Webster in his article, "America's Highest Ranked Graduate Schools" (1983) reviewed the above Assessment study and noted two serious weaknesses. First, a number of the variables used did not relate to quality, and, secondly, no attempt was made to rank the schools or programs--only alphabetical lists of schools were presented. Webster, using the Assessment data, goes on to rank the schools and programs according to reputational quality of their graduate faculty. In addition, he develops a method for ranking graduate schools which allows for analysis of the rise and fall of schools over those years by using the results of seven major reputational studies covering the 1925-1982 years.

Webster's data can be helpful when used with the RAI. The school rankings based upon the Assessment data can be compared to the R&D expenditure rankings and the RAI rankings to gain an understanding of how quality relates to research dollars and how quality relates to research activity/scale.

The RAI differs from the above studies in that it has no judgmental variables and only uses objective, descriptive variables. The RAI does not attempt to measure quality but measures research activity and scale of research programs. There may be a measure of quality in the RAI and the RAI may be an appropriate measure for input into future quality studies, but that is not the primary purpose for developing the RAI.

The Assessment study provides data on sixteen different variables. The reader is forced to evaluate each variable before attempting to combine the variables into some overall comprehensive ranking for a particular institution. By comparison the RAI is a composite index of many variables. The RAI is designed to assist the reader in evaluating the ranking of an institution.

Even though the RAI is based upon descriptive data, some of Arrowsmith's objections may apply. It is true that the "overall effect" of ranking research institutions may make the "rich, richer" and the "poor, poorer."; however, that is dependent upon the philosophy of university funders.

If egalitarianism in funding is desired then the opposite effect could be obtained. Arrowsmith indicates that there is a tendency to reduce diversity and to reward conformity which "... might be tolerable if the reports were tolerably accurate and judicious, if they were less prescriptive and more descriptive..." (Dolan 1976, p. ix).

The RAI excludes the prescriptive judgmental variables that Arrowsmith refers to and uses descriptive variables. Another criticism of the reputational ratings is that they "... inherently reflect perceptions that may be several years out of date ..." (Jones et al., eds., 1982, p. 5). The reasons such studies are out of date are (1) they are based upon faculty perceptions that are developed over time, and (2) national surveys are very costly and have only been performed every decade or so. The RAI can use variables that are a year or two old and are readily available. Therefore, RAI's can be generated annually and would reflect current reality.

Other criticisms are that ratings receive a halo effect from the overall reputation of a university, and a disproportionate number of evaluators are from the larger universities. The RAI would eliminate halo and alumni biases because it would be based upon objective, not judgmental data.

Examining the Assessment study variables, provided information that was useful in selecting the variables that went into developing the RAI. Of the sixteen variables used in the study, three were also used in the RAI: the number of faculty members, the ARL index, and research & development expenditures. The other thirteen <u>Assessment</u> variables were the results of special one-time surveys and did not meet the availabilty criterion of the ARL index or were not directly related to research activity. The program size variables, which included numbers of faculty, students, and graduates for the past five years, are not related to quality per se but are "...thought to have a significant influence on the effectiveness of programs" (Jones et al., eds., 1982, p. 16). It was noted also that "in previous studies program size has been shown to be highly correlated with reputational ratings of a program..." (Jones et al., eds., 1982, p. 16). The ARL, being a measure of research activity and size, may likewise be highly correlated with reputational ratings.

The <u>Assessment</u> study took several approaches to determine the validity of the variables used in its study. These approaches have equal applicability for validating RAI input variables. According to the study, "Measures that are logically related to program quality are expected to be related to each other. To the extent that they are related, a stronger case might be made for the validity of each as a quality measure." (Jones et al., eds., 1982, p. 163). The Pearson product-moment correlation coefficients were examined and strong correlations were found between key variables. Also, consideration was given to providing an alternative set of variables adjusted for program size, but a satisfactory algorithm was not found. It was found that exceptionally large programs appeared to be unfairly penalized, and the reverse occured for some smaller programs (Jones et al., eds., 1982, p. 165). Finally, the reputational ratings were compared to the ratings of the prior Roose-Anderson study because if they were comparable then they would tend to validate each other. A high degree of correlation between the two studies was found.

The correlation coefficients of the variables used in developing the RAI can be examined, and they would be expected to be high if they are in fact related to research activity. The original proposal for the RAI indicated that adjustments for size would be made. This was abandoned because it was believed that size was an important dimension in the RAI. If the experience of the Assessment study is applied to the RAI, it may indicate that if size had been controlled in the RAI some distortions may have been experienced in the cases of the very large and smaller institutions. The procedure for comparing a new index with an older index can be adapted to the RAI study. The RAI can be compared to the R&D expenditures ranking.

Related Studies

The 1981 study by McCoy, Krakower and Makowski, Financing of the Leading 100 Research Universities: A Study of Financial Dependency, Concentration and Related Institutional Characteristics, approximated the sample of institutions and employed a number of the variables that were used in developing the RAI. The study examined financing trends, the extent to which institutions are dependent upon federal funding, the concentration of federal funding among all research providers and higher education research providers, and the relationships between research funding and specific characteristics of research institutions. The data used were from the NSF Science Resources Studies program, NCES, and HEGIS surveys.

The McCoy et al. study found that leading research universities receive significant shares of their funds from the federal government, and high rates of government funding are accompanied by substantial funding from non-federal sources. Also, since 1960 federal R&D funds for higher education have increased at the expense of industry based research. No major shift in federal R&D funds was noted within higher education between 1975 and 1979. A linkage between research and graduate, primarily PhD-level, education was noted. The study also found that land-grant status did not appear to be an important variable. Among the data-related findings, it was found that the NSF and

NCES data bases could be used in parallel. Also, a number of institutions had to be discarded from the study either because they reported system-wide data rather than campus based data or because of missing or inadequate data (1981, pp. 195-319d).

A proposition used in developing the RAI was that a univariate indicator i.e. R&D expenditures, may not be stable enough from year to year to provide a reliable basis for rank-ordering research universities. Examination of R&D expenditures for 200 universities for the 1976 through 1979 years reveals considerable variation for some institutions (McCoy et al., 1981, Appendix-A, pp. 319a-d). An example of the fluctuation in funding is shown in the following table where institutional changes in dependency on federal contracts and grants for support of R&D are presented:

Table 2.1. Changes in Dependency on Federal Funds for Support of R&D*

Ranges of Change in Percentage Points 1975-1979 NSF Data	Total Institutions with Changes	Percent Of Total
Over -6%	11	(15.5%)
-3.5 to -6	11	(15.5%)
-1.5 to -3.4	10	(14.1%)
-1.5 to 1.5	18	(25.4%)
1.6 to 3.4	7	(9.9%)
3.5 to 6	6	(8.5%)
Over 6	8	(11.3%)

* Source: McCoy et al., 1981, Table 2-18, p. 75.

In the five years covered, 15.5 percent of the institutions had over a -6 percent change and 11.4 percent had over a +6 percent change in R & D funding. The above data indicate that a ranking based upon financial data alone may be unstable for some institutions. An index should be relatively stable and not overly sensitive to short-term fluctuations. By combining a number of variables into the RAI, fluctuations or data errors in any single variable could be smoothed and would better represent an institution's relative position than a univariate index.

Support for the decision to limit the RAI to the top 100 or so research universities is found in the McCoy study. The leading research institutions conduct almost 85 percent of the R&D in higher education and produce 40 percent of all graduate degrees in science and engineering. Their average revenues and expenditures are more than 10 times greater than institutions in the Other Graduate Institutions category. After adjusting for differences in faculty size and enrollments, R&D expenditures in the top 100 institutions, compared to the Other Institutions, are more than 5 times higher in the public sector and 14 times higher in the private sector. It is clear that the leading research institutions are larger, concentrate more on graduate programs, and are better financed than other institutions (McCoy et al., 1981, pp. 125-129).

A number of multiple correlations in the McCoy study help support the selection of particular variables for use in developing the RAI. The correlation between total full-time faculty and R&D expenditures (.65) is greater than the correlation between total FTE student (graduate and undergraduate) enrollment and R&D expenditures (.56), indicating that institutional R&D efforts may be more influenced by the total number of full-time faculty than by student enrollments (1981, p. 145). The correlation between graduate FTE students and R&D expenditures is .76, while the respective correlation with undergraduate students is only .44. The correlations between R&D expenditures and the five-year average degrees awarded are .64 for BA degrees, .77 for MA degrees and .89 for PhD degrees. The relatively high correlations of R&D expenditures with graduate FTE students and degrees awarded reinforce the validity of selecting these variables for inclusion in the RAI.

David J. Bowering (1981) performed a major study entitled, <u>Draft Report on the Impacts of Federal R&D Expenditures on Graduate Student Enrollments, Graduate Degree Production, and Professional Staff Size at Leading Research Universities</u>. Three fields of science were examined: physical science, biological science, and engineering. Bowering maintains that many studies involving enumeration or correlation imply causation or leave it up to the reader to make such implications without justification for doing so. His study attempts to separate the causal from the non-causal elements using path analysis and causal modeling techniques. The findings, which Bowering notes should be considered as hypotheses in future studies, show that over the 1975-1979 period federal R&D expenditures played a significant role in attracting and retaining PhD students, in the growth and maintenance of professional R&D staff, and in facilitating PhD completions (1981, pp. 2-22).

If causal relationships are established between variables in the Bowering study and the same variables are used in the development of the RAI, then additional credence is lent to the validity of the RAI variables. If one accepts the Bowering study then research funding causes changes in graduate student enrollments, graduate student degree production, and professional staff size, all of which represent elements of research activity. The R&D expenditures, the professional staff, and the graduate student variables were from the NSF survey data bases, which are the same source used for the RAI. Data for the number of PhD's awarded came from a different source, but it is doubtful that significant variations in the data would exist because both were furnished by the same institutions.

Edgar J. Ellyson and Jack P. Kruegar in their 1980 study, Predicting Federal Research Funding at Colleges and Universities, examined the relationships between federal research funds and institutional research funds, degrees awarded by level, and student headcounts by level. NSF survey data from some 60 research universities were subjected to multiple regression and correlation analysis. The results showed that Phd degrees awarded and institutional research funds had the strongest correlations with federal research funding (1980, P. 131). Ellyson's table of correlations can be compared to the correlations generated in the RAI study as a check on the reliability of the RAI data.

David A. Katz in The Impact of Federal Grants on Output and Employment in Universities (1980) determined that federal grants had a significant impact on output and employment of university professors. This finding lends support to using citation indices and number of faculty as two variables in developing the RAI.

David E. Drew in his 1975 paper Science Development: An Evaluation Study examined the changes in the quality of graduate education and research productivity in three fields of science at institutions that participated in the NSF Science Development program. The following variables were used in the study: graduate enrollment, doctoral production, number of science faculty, and federal funding. The use of these variables in this study lends support to their use in the RAI.

While many ratings have been controversial, David R. Gerhan in his article, Graduate & Professional School Ratings: Facing the Need to Know (1979, p. 215) argues that there is an important demand by graduate school applicants, employers, fund granting agencies, and the general public for such ratings. Gerhan's study was to discover, collect, and report as many recent rankings and comparative studies as possible. His findings note that "The 'top ten' variety does indeed exist. Nevertheless, another pattern also appears: a subset of empirically based, nonevaluative studies which permit limited, fairly objective comparisons among schools" (1979, p.217). A bibliography that includes 74 ratings or comparisons done since 1972 is presented. Two thirds (50 out of 74) of the ratings were comparisons based upon descriptive, objective criteria.

Summary

A clear distinction was drawn between the subjective, judgmental type quality indices used to evaluate universities and the objective, descriptive type indices like the RAI. It was also noted that there is some instability in R&D funding from year to year for a number of institutions, thus bringing into question their reliability. This underscores the importance of using multiple measures in evaluating universities if reliability is to be achieved. In terms of the variables selected for developing the RAI, a number of the variables have a causative and/or correlative relationship to R&D funding, thus supporting their selection for inclusion into the study.

CHAPTER 3

DATA SOURCES AND RESEARCH DESIGN

This chapter is divided into five sections. The first section reviews the research questions to be tested; the second section examines the analytical framework for the development of the RAI; the third section presents the data sources; the fourth section outlines the statistical methods; and the final section presents some special limitations of the study.

Research Questions

There were two research questions: Can an index be developed that will provide a composite indicator of research activity at the major research universities? Does such a index provide a ranking of universities that is more valid and reliable than existing univariate indicators? The first research question was tested using statistical procedures, and the second was subjected to normative quantitative analysis.

Analytical Framework

The analytical framework for the derivation of the RAI was based upon the work of Kendon Stubbs (1980), who describes the development of an index for the Association of Research Libraries (ARL). Stubb's work can be described as a four step process in which (1) data were gathered on a number of research library and university variables; (2) through the application of statistical procedures, a related subset of variables was identified; (3) a statistical technique called principal component analysis was applied to the subset of variables to generate a weight or loading factor for each variable; and (4) an index score for each library was created by multiplying each variable's component loading (coefficient) by the library's data values and then summing the results to form the ARL score.

Specifically, as a first step, Stubb's method (1980, pp. 1-12) of developing the ARL index was to use data from the ARL annual survey of member libraries. From twenty-five variables which included such items as total library expenditures, salary expenditures, number of personnel and number of volumns, Stubbs used factor analysis and identified four different factors. The first was a library size and resources factor; the second and third factors were related to inter-library borrowings, and the fourth factor was related to university size. The first factor had fifteen variables that were loaded heavily. Within these fifteen variables, some were subsets or combinations of others and were dropped from the study. The student assistants variable loaded equally on factors one and four; however, both relations were weak so the variable was dropped. The remaining ten variables were used in the subsequent steps. In the third step, a variant of factor analysis called principal components analysis was applied to the first factor. Component score coefficients were generated for each of the ten variables. These coefficients were multiplied by the standardized data values for each library, and the linear combination of these products became the ARL index score for each library.

A normalized score of -1.75 was required to permit continuing membership in the ARL. The index scores were also used to determine whether a new library should be admitted into the ARL. A score of at least -1.00 was required for new members. The ARL index has been used since 1980 and is published annually in The Chronicle of Higher Education. The index was used in the Associated Research Council's study, An Assessment of Research-Doctorate Programs in the United States (Jones et al., 1982, p. 15). These uses of the ARL index lend support to the acceptability of the ARL method.

Data Sources

In this section the population of research universities is presented, along with specifications criteria for selection of the variables and primary data sources. A table of research variables and their descriptions is presented.

Population

An initial population of the 112 research organizations with the largest 1980 research and development expenditures was selected. A number of research organizations had to be excluded because of either incomplete data or data that included more than one campus. A total of 112 organizations was initially included and 68 survived the final reductions (see Table 3.1). These reductions were consistant with the experience of other studies using these research universities (McCoy et al., 1981; Ellyson and Krueger, 1980).

Table 3.1. Population of Universities-Ranked According to R&D Expenditures

FICE	Universities
2077	Johns Hopkins
2178	Mass. Inst of Technology
3895	Univ. of Wisconsin-Madison
1317	Univ. of Cal.-San Diego
1305	Stanford Univ.
3798	Univ. of Washington
2707	Columbia Univ.-Main Division
2155	Harvard Univ.
3378	Univ. of Pennsylvania
1312	Univ. of Cal.-Berkeley
1315	Univ. of Cal.-Los Angeles
1775	Univ. of Ill.-Urbana
3658	Univ. of Texas at Austin
1328	Univ. of Southern Cal.
1319	Univ. of Cal.-San Francisco
2290	Michigan State Univ.
1426	Yale Univ.
1083	Univ. of Arizona
1313	Univ. of Cal.-Davis
2894	Univ. of Rochester
8732	Purdue Univ.
1774	Univ. of Chicago
2785	New York Univ.
1535	Univ. of Florida
1598	Univ. of Georgia
1869	Iowa St. U of Sci & Tech.
3210	Oregon State Univ.
1131	California Inst. of Tech.
2972	N C State Univ. at Raleigh
2903	Yeshiva Univ.
1739	Northwestern Univ.
1536	Univ. of Miami
3024	Case Western Reserve Univ.
1350	Colorado State Univ.
1610	Univ. of Hawaii-Manoa
2103	Univ. of Maryland-College Pk.
2920	Duke Univ.
2974	Univ. of NC at Chapel Hill
1892	Univ. of Iowa
2807	Rockefeller Univ.
3675	Univ. of Utah
2516	Univ. of Missouri-Columbia
3754	Virginia Polytech Inst&St. U.
2565	Univ. of Nebraska-Lincoln
3800	Washington State Univ.

Table 3.1--Continued.

FICE	Universities
1052	Univ. Alabama-Birmingham
1809	Indiana Univ.-Bloomington
3242	Carnegie-Mellon Univ.
2423	Mississippi State Univ.
2627	Princeton Univ.
1928	Kansas St. U-AG & App. Sci.
1316	Univ. of Cal-Riverside
3677	Utah State Univ.
2130	Boston Univ.
1314	Univ of Cal.-Irvine
1489	Florida State Univ.
2329	Wayne State Univ.
3535	Vanderbilt Univ.
3127	Univ. of Dayton
1444	George Washington Univ.
3371	Temple Univ.
3735	Virginia Commonwealth Univ.
3401	Brown Univ.
3827	West Virginia Univ.
1445	Georgetown Univ.
1431	Univ. of Deleware
2221	Univ. of Mass at Amherst
8789	Syracuse Univ.

Criteria

Victor Went (1980, pp. 22-23), of the National Center for Educational Statistics, in reviewing the federal interest in indicators for federal policy purposes, indicates that the following characteristics of indicators are needed:

> 1. a conceptual basis readily understandable by all potential audiences;
> 2. clear delineation of the proper uses and limitations of the method used in building the indicator, its results, and its possible misuses;
> 3. ease of gathering and updating reliable and timely data.

Characteristics such as these appear useful and items 2 & 3 were incorporated into this study. A conceptual basis using input/output analysis theory is presented later in this study.

In addition to Went's suggestions, the criteria for the selection of variables evolved in two ways. First, variables had to possess face validity; they had to be reliable, that is, stable over time; and they had to be readily available. Secondly, formative development of criteria was accomplished through statistical analysis that allowed for the elimination of unimportant variables. Specifically, the list of variables was pruned through correlation analysis by elimination of variables with low multiple correlations. Also, the surviving variables had to make a substantive contribution to the RAI through their component loadings in the principal component analysis.[1]

Primary Data Sources

A preliminary review of data sources included four National Science Foundation longitudinal data bases and the Institute for Science Information citation indices. Also reviewed were the National Center for Higher Education Management Systems data; percentage of federal grants to total grants, index of change in federal grants, index of change in education and general expenditures, percentage change in federal dependency ratios for the top 100 research universities, change in federal research and development dependency ratios, and various indices of change in federal and total expenditures. Likewise reviewed were the National Center for Educational Statistics data, the percent of tenured faculty by rank, and the percent of graduate

----

1. Recognizing the importance of this topic to institutions and the controversy that may arise, Dr. Chester W. Harris, who is a recognized expert in factor analysis, served as a statistical consultant for this study. Dr. Harris is past president of the Psychometrics Society, the Society of Multivariate Experimental Psychology, and is widely published in the field.

students on federal support. Astin' index of selectivity scores, Halstead's Research and Development Price Index, the Gorman rankings, and various other quality indices were also reviewed.

Of the potential sources, the following passed the criteria established for the RAI:

1. National Science Foundation. NSF Survey of Scientific and Engineering Expenditures at Universities and Colleges. Washington: Government Printing Office, 1980. This is one of four annual NSF surveys that were available. Although data were available for ten years, only 1980 data were used in this study. This survey provided the financial data for the study. Research has indicated that the major alternative financial data base, the HEGIS surveys, may be unreliable when used for institutional level analysis (Minter, 1979, pp. 16-18).

2. National Science Foundation. NSF-NIH Survey of Graduate Science Student Support and Postdoctorals. Washington: Government Printing Office, 1980. This survey provided graduate science student enrollment data.

3. National Science Foundation. Survey of Scientific and Engineering Personnel Employed at Universities and Colleges. Washington: Government Printing Office, 1980.

This survey provided graduate level personnel data in the sciences.

4. American Association of University Professors. <u>Rocky Road Through the 1980's: Annual Report on the Economic State of the Profession</u>. Academe. Washington: AAUP, 1980. This source provided faculty salary data.

5. Stubbs, Kendon. <u>The ARL Library Index and Qualitative Relationships in the ARL</u>. Washington: ARL, 1980. This index was derived from ten variables that provided a measure of research library size.

6. National Center for Educational Statistics. <u>Higher Education General Information Survey</u>. Washington: NCES, 1980. The educational and general expenditures variable was obtained from this data base.

Variables

A large number of variables were selected and screened for inclusion into this study, but most failed to pass the criteria for acceptance. The variables were selected by reviewing the relevant literature and examining the existent data bases. A number of studies have used variables that were related to research activity: Bowering 1981, Drew 1975, Ellyson and Krueger 1980, and McCoy, et al., 1981. Table 3.2 lists the variables that passed the first cut and were used in this study.

Table 3.2. Research Variables

Name	Source	Description
AVGEQP	NSF*	Four year average equipment expenditures (1).
ENDEB	HEGIS	Endowment value, end of year, at book value (1).
ENDEM	HEGIS	Endowment value, end of year, at market value (1).
PCTCHG	NSF**	Average percentage change in R&D expenditures from 1977-1980 (computed).
VAR1	NSF	Source of Funds-Federal government (a subset of source of funds total) (5).
VAR3	NSF	Source of funds-state and local governments (a subset of source of funds total) (1&5).
VAR5	NSF	Source of funds-industry (a subset of source of funds total) (3&5).
VAR7	NSF	Source of funds-institution's funds (a subset of source of fund s total) (5).
VAR9	NSF	Source of funds-all other sources (a subset of source of fund s total) (5).
VAR11	NSF	Source of funds total. Same as total R&D expenditures.
VAR193	NSF	Total capital expenditures for scientific and engineering facilities and research equipment for research, development and instruction. Estimates were provided for two institutions with missing data.
VAR195	NSF	Federal capital expenditures for research facilities and equipment (a subset of var193) (3).
VAR197	NSF	All other-capital expenditures (a subset of VAR193) (3).

Table 3.2--Continued

Name	Source	Description
INST1	NSF	Full-time scientists and engineers (1).
INST3	NSF	Part-time scientists and engineers (1).
INST5	NSF	Full-time equivalent scientists and engineers (5).
SINST1	NSF	Full-time science graduate students.
SINST2	NSF	Part-time science graduate students.
SINST3	NSF	Post-doctorates
SINST4	NSF	Other non-faculty science doctoral Research staff.
SINST5	NSF	Post-doctorate and other science students (2).
EG	HEGIS	Education and general expenditures and mandatory transfers. Two institutions required estimates for missing data (4).
FCOMP	ARL	Average faculty compensation. Three institutions required estimated data(1).
PHDS	ARL CEEB	PhD's awarded for 1980.
FLDS	ARL	Academic fields of study (1&3).
ARLIND	ARL	The Association of Research Libraries composite index. Fourteen institutions did not have ARL index numbers, these were estimated at the lowest ARL value (-1.56) since none of these institutions qualified for the ARL.

(1). Variable was eliminated because correlations or factor loadings were small.
(2) Variable was eliminated because it summarized other variables.
(3) Variable was eliminated because of excessive missing data values.

Table 3.2--Continued

(4) The E&G Expenditure variable was eliminated because the research activity relationship was not a direct one (i.e. the departmental research that was included in E&G expenditures could not be separately identified).

(5) Variable was eliminated because it was part of a redundant subset of a total variable and did not add information to the RAI.

*The AVGEQP variable represents the average facilities and equipment expenditures for the years 1976-80.

** The PCTCHG variable, which is the average percentage change in research and development expenditures from 1976-80, was computed as follows: each year's research and development expenditures were divided into the following year's expenditures to arrive at the percentage change for each year. These were summed for the four years then divided by four years to arrive at the average percentage change.

Method

Research Question Number One

To answer the first research question an index was developed in four stages: (1) a data base of research related variables was created; (2) the data were reduced; (3) the principal components were selected; and (4) the research activity scores were computed.

Stage one required that a number of Fortran computer language programs be developed to extract the selected institutions and variables from the three National Science Foundation data bases and the National Center for Educational Statistics data base. Several variables were computed using multi-year data. Also, a separate file of eight variables from the Association of Research Libraries and the College Entrance Examination Board was created and merged with the selected variables from the other data bases. The result was one file with all of the selected variables for each institution.

A number of problems necessitated discarding various institutions from the study. Frequently, institutions reported system-wide rather than institutional data or they had too many missing data values. The details related to the selection of the population of institutions are presented in Appendix A and the details associated with the missing data values are presented in Appendix B.

Stage two, the data reduction stage, involved two steps: (1) examination of the multiple correlation coefficient matrix for variables with weak correlations, and (2) examination of the principal component score coefficients of each variable to determine which variables were not contributing significantly to the principal components. Pearson product-moment correlations were generated using the <u>Statistical Package for the Social Sciences</u> (Nie et al., 1975) on the initial twenty-six variables. Those variables that had consistantly weak multiple correlations were excluded from further analysis; those that had low correlations with a number of other variables but had strong correlations with at least three variables were included in the following analysis. A principal components analysis was generated with the expressed purpose of examining the score coefficients of each of the remaining variables to determine the magnitude of their relative contributions to the principal components. If a variable had either a large positive or negative principal component loading and its correlation coefficients were large on several variables, the variable was retained in the study. However, if the principal component loading approached zero and the correlation coefficients were only strong on a few variables, then the variable was excluded from further study. Here principal component analysis was used as a data reduction technique, as contrasted to its later use as a descriptive technique to build the index.

It is interesting to note that Stubbs in the Association of Research Libraries study (1980, pp. 8-10) used factor analysis for data reduction. Factor analysis was not selected for this study because, in the interest of conservatism, the introduction of factor analysis with its attendent inferential assumptions (i.e. using estimates in the correlation matrix diagonal) may confuse the straightforward principal component method. While data reduction was of some concern, it was not a major concern because the data were readily available.

In stage three, principal component analysis is used as a descriptive statistic. It is a statistical technique for transforming linearly a set of variables into new sets of composite variables or principal components. Principal components are uncorrelated (orthogonal to each other). For any such component, each variable in a principal component analysis has a component loading which may be standardized and multiplied by the corresponding standardized variable data values for each institution in the study. The resulting scores for each variable may be summed into a principal component score for the institution which is called the Research Activity Index. The scores may be sorted in descending order to rank the institutions.

Principal component analysis was selected because (1) it provided a statistically optimal solution in terms of explaining as much variance in the data as possible (Harmon 1976, p. 4; Rummel 1970, p. 112; Daultry 1976, p. 3), (2) principal component scores for each institution were determined exactly and with no inference or estimation (Harmon 1976, p p. 365-66); and (3) the principal component scores provided a means to rank order institutions (Daultry 1976, pp. 44-5).

Based primarily upon Daltrey's work, principal component analysis specifically can be described as a data transformation technique applied to a number of cases with multiple variables that generate a first principal component accounting for as much of the total variance in the data as possible. The second principal component accounts for as much of the remaining variance as possible while remaining uncorrelated to the first principal component. Additional principal components account for the remaining total variance and are also uncorrelated to any other principal components. The principal components can be interpreted in terms of the relative contributions (loadings) of their variables. Each variable has a correlation with the principal component; this value is referred to as the variables component loading. Squaring each variable's component loading produces the proportion of the variance of each variable that is explained by each component (1976, p. 22). To generate an index for an institution, the component loading for each variable was multiplied by the standardized data values for each institution to arrive at a score for each variable. The linear combination of these scores became the score for each university. "All that is being done in principal component analysis is the transformation of the system of reference from one set of axes (the variables) to another (the principal components). Nothing 'new' is created; nothing 'old' is lost" (Daultrey 1976, p. 23). The reader is referred to several references for technical explanations of principal component analysis. Daultrey (1976) explains and applies the techniques. Also, Harmon (1976); Kim and Mueller (1978); Nie, Hull, Jenkins, Steinbrenner, and Bent (1975); Rummel (1970); and Velicer (1976) elaborate on principal component analysis.

A primary issue to be resolved is how many components will be examined, considered important and interpreted. If the maximum variance in the variables is to be accounted for, then the first principal component should become the basis for producing the RAI. However, if a more definitive interpretation of the components is important, then the components can be rotated, criteria can be established to determine which components are significant, and the components can be interpreted. Henceforth, in this paper these two methods will be refered to as the first principal component and the rotated methods. Each method can result in a different ranking of the institutions. The theoretical aspects of these methods will be presented in the following paragraphs and the empirical findings will be presented in Chapter 4.

In principal component analysis, often considered as a branch of factor analysis, a correlation matrix of all the variables is generated with unities in the main diagonal. The statistical technique extracts one or more principal components which are exact linear transformations of the original variables. These exact mathematical transformations are the reason for selecting principal component analysis for developing the RAI because they do not require any inferential assumptions about the general structure of the variables. The first principal component is the best single summary of the linear relationships in the data (Kim et al., 1978, p. 470). The first principal component accounts for more variance in the data than any other solution. This is appropriate when care is exercised in the selection of variables so that there is a balanced group of research related variables that go into the principal components analysis (Kim et al., p. 482). These are strong theoretical arguments for using the first principal component in the development of the RAI.

Rotation can be used to change the component structure so that each component only loads on a few different variables. This is called the Thurstone simple structure notion. If it is true, then variables can be sorted into distinct sets of variables and interpretation of the components is facilitated. They are independent clusters which can be correlated or uncorrelated to each other. If it is not true, then some of the variables are of a complexity greater than one (Harris, 1983). A second reason for using a rotated solution is that loadings in first principal component solutions depend heavily upon the relative number of variables. For example, the deletion of one variable which stands for a certain theoretical dimension could change drastically the relative loadings on the first principal components. The rotated components are more stable in this respect than the first principal components (Nie et al., 1978, p. 483). Dr. Chester Harris in reviewing this statement argues that this is not true (Harris 1983). Nevertheless, there are good theoretical reasons for rotation of the principal components.

When a rotated solution is desired there are a number of different rotation options available. The option that was selected for this study is called VARIMAX. This option centers upon simplifying the columns or components. Varimax maximizes the variance of the squared loadings in each component.

In recapping the first principal component versus rotated issue, the first principal component solution will maximize the variance accounted for on the initial principal component when the variables are carefully selected.

However, the rotated solution may allow for the selection and identification of a component and therefore may provide a better conceptual basis for the RAI. The rotated RAI may also be more stable in the event future first principal component RAIs have particular variables added or removed. Since there is no clear-cut theoretical preference, both solutions will be presented throughout this study. They will be evaluated by comparing and contrasting each RAI to each other and to the R&D expenditures ranking.

Another question to be addressed relates to the combination of a number of principal components to generate an RAI. The rationale is that the RAI should account for as much variance in the data as possible. While the first principal component accounts for the most variance, the remaining components account for additional variance that could be incorporated into an RAI. Several problems must be resolved to accomplish combining components into one RAI. These reflect how many components should be included and how the component scores can be weighted and summed into one RAI score. The components to be included can be those with an eigenvalue greater than one. This is an accepted rule of thumb that has the effect of excluding minor, unarticulated components from further consideration. The components can be weighted according to their proportion of total variance. Specifically, each component's variance (eigenvalue) is divided by the total variance of all the included (eigenvalues greater than one) components to arrive at the component's percentage of total variance. This percentage of total variance is multiplied by the components score to arrive at the weighted score for each component. These weighted component scores are summed to arrive at a combined RAI score. The formula for this method is as follows:

$$\mathrm{RAI} = \sum_{i=1}^{m} (C_i) \left(\frac{\lambda_i^{m}}{\sum^{m} \lambda_i} \right)$$

Where m = the components with an eigenvalue greater than one, λ_1 = the eigenvalue (amount of total variance accounted for by component one) and C = the composite scale (scores) for each component.

The components with an eigenvalue greater than one will be combined for both of the RAIs. The resulting rankings will be compared to those generated from the first principal component and the first rotated component.

While some of the strengths of principal component analysis have been outlined, it is important to examine some of caveats associated with the technique. Daultrey (1976, p. 41) points out that measurement error can cause biased results. Such error can be introduced by using estimates for missing data values, using inconsistant data definitions and categories, and by mistakes in compiling data. Measurement errors when systematically biased (correlated) are not easy to correct. However, random errors can be accomodated since the effect is to lower the expected magnitude of correlations among the variables (Kim and Mueller 1978, p. 68). When selecting variables, a selection bias can predetermine the results. Daultrey (1976, p. 43) indicates how he reduced bias by eliminating three of six variables that were measures of the same characteristic. The reduction was a decision based upon knowledge of the data. By exercising care in selecting variables, particularly in eliminating redundant variables and variables that are subsets of other variables, selection bias can be minimized.

Stage four involved the calculation of the principal component scores for each university. This is an option in the SPSS program, and it produces a single index score for each university. These scores were sorted and listed in decending order to generate the Research Activity Index ranking.

Research Question Number Two

Research question number two asks whether the RAI provides a ranking of universities that is more valid and reliable than any existing univariant measure? The RAI was validated by checking to insure that no particular set of variables dominated or caused a bias in the results, by analysis of each variable's contribution to the outcome, by comparing the index with the R&D expenditures ranking, and by having the statistical methodology reviewed by a principal components analysis expert. The reliability of the index is best tested by developing it over a number of years; however, this was beyond the scope of this study. Reliability was partially established by reviewing the variables stability in previous studies, by examination of the variables for stability and consistancy over time, and by determining that like institutions produce like RAI scores.

Special Limitations

There are several limitations that are related to data availability. The first is that about twenty-five percent of the top 100 research universities report their data on a system-wide bases rather than on a individual institution basis. A second limitation is that there were excessive missing data for a number of universities.

As previously noted, the National Science Foundation surveys do not collect data on all fields of study. These exclusions affect the comprehensiveness of the study. However, since the total dollars involved in research and development in these excluded fields is relatively small the impact of this limitation should not seriously affect the rankings.

Summary

The two tasks of developing the Research Activity Index and testing the hypotheses were outlined in detail. The university sample was identified; the criteria for selection of the research variables were established; and the research variables were identified. The statistical method followed a four-stage process involving the development of the data base, data reduction, principal component analysis, and production of the Research Activity Index scores. Finally, the question of determination of the validity and reliability of the index was examined, as were the limitations of the study.

CHAPTER 4

FINDINGS

This chapter presents the findings as to whether a composite Research Activity Index can be developed. Also, the question of whether such an index would be a more valid and reliable measure than existing univariate indicators is addressed. Specifically, the findings associated with developing the RAI were related to data reduction, selection of the best principal component method, and computation of the RAI scores. The findings associated with validity involved interpretation of the components, examination of the variables impact upon the RAI, analysis of the institution's ranks for reasonableness, and a review by a principal components analysis expert. The reliability findings involved looking at the consistency of the variable definitions over time and a determination that similar RAI results were produced by similar institutions.

Development of the RAI

Data Reduction

A number of issues surround the selection of variables for the RAIs. These are (1) what variables, as evidenced by low multiple correlations and/or component loadings, do not contribute significantly to the RAI scores;

(2) should variables that are subsets of other total variables be discarded; and (3) are related clusters of variables, such as, financial, personnel or student variables, too heavily represented in the RAI, thus overloading or biasing the RAI in a particular direction? Five computer runs were executed, with analysis and evaluation of the results after each run. Each was followed by a successive elimination of variables that did not contribute significantly to the component scores.

Starting with 26 variables (See Table 3.2), the first variable eliminated was the educational and general expenditure variable, which included unspecified departmental research. The variable was eliminated because such total expenditure was, on the face of it, not directly related to research activity.

Examination of the correlation coefficients showed that the following variables generally had weak correlations with the other variables: ENDEB (endowment value, end of year, at book value), ENDEM (endowment value, end of year, at market value), AVGEQP (four year average equipment expenditures), VAR3 (source of funds- state and local governments) and FCOMP (average faculty compensation). Examination of the component loadings indicated that the following variables did not have significant loadings on the components (eigenvalues greater than one): ENDEB, ENDEM, AVGEQP, FCOMP and part-time students. These variables, which were weak on both the correlation matrix and on the principal component loadings, were eliminated. Those variables that were weak on either the correlation matrix or the principal component loadings but not on both were subjected to further analysis.

Fifteen institutions had missing values in the FLDS (academic fields of study) variable. Another weakness in the FLDS variable was that it represented both undergraduate and graduate fields of study, whereas research is primarily associated with graduate study. Also, the FLDS correlations with other variables were weak. For these reasons the FLDS variable was eliminated.

Next, a determination was made as to whether each variable was redundant and therefore introduced double counting or weighting. However, several statisticians advised that elimination of subset variables would result in loss of discriminating information. Therefore, rather than lose the diverse information in the subset variables, it was decided to exclude initially the overall, or total, variables. These were R&D expenditures, capital expenditures and SINST5 (postdoctorate and other students). Likewise, INST5 (full-time equalivent scientist and engineers), was to a certain extent, redundant of the full-time and part-time scientists and engineers (headcount) variables and therefore was eliminated.

Upon examination of the component weightings, it was observed that some institutions were receiving lower scores for having large sources of funds from atypical sources. Since the rationale for using the financial variables was that they provided a proxy for research activity, it followed that the particular source of funds was irrelevant and that only the total source of funds variables were needed.

As an illustration, due to large differences in its sources of funds, the University of California at Berkeley showed a rotated component score coefficient of -.22 for VAR7 (source of funds-university). The mean source of funds for all institutions was $6,388 and Berkeley's figure was $18,405. The calculation of the score for this variable resulted in a negative score for Berkeley. The formula for standardizing the variable is Z =(VAR7-mean of VAR7)/standard deviation of VAR7, (($18,405-$6,388) / $6652=1.806). The rotated component score for all universities was -.12. Berkeley's component score for VAR7 would have been (-.12x1.8060)=-.22. (This -.22 value is combined with values for each variable to arrive at the RAI score.) The question was why should Berkeley's RAI score be reduced because it happened to have sources of funds different from other schools? It was decided that the particular sources of funds were not important, but the magnitude of total funds was. Therefore, VAR1, VAR3, VAR5, VAR7, VAR9, (source of funds subsets) were replaced by R&D expenditures, the total source of funds. Likewise, VAR195 and Var197, capital expenditures subtotals, were replaced by total capital expenditures. Table 4.1 shows the final eleven variables that were used in the remainder of this study.

A review of the final eleven variables to determine whether any group of related variables was overly represented indicated that there were three financial variables, two personnel variables, five student variables, and one library variable. Part-time students had a small component weighting and therefore had little impact on the RAI values. These student variables were reevaluated during subsequent analysis to make sure that they did not dominate the RAI.

TABLE 4.1. Final Research Variables

Name	Source	Description
percentage change	NSF	Average percentage change in R&D expenditures from 1977-80 (computed).
R&D expenditures	NSF	Source of funds total. Same as total R&D expenditures.
capital expenditures	NSF	Total capital expenditures for scientific and engineering facilities and research equipment for research, development and instruction. Estimates were provided for two institutions with missing data.
full-time scientists	NSF	Full-time scientists and engineers employed.
part-time scientists	NSF	Part-time scientists and engineers employed.
full-time students	NSF	Full-time graduate science students enrolled.
part-time students	NSF	Part-time graduate science students enrolled.
post doctoral	NSF	Post-doctoral study
other students	NSF	Other non-faculty doctoral research staff/students.
PhD awards	ARL CEEB	PhD's awarded for 1980.
ARL index	ARL	The Association of Research Libraries composite index.*

* Fourteen institutions did not have ARL index numbers: these were estimated at the lowest ARL value of -1.56 since none of these institutions qualified for the ARL.

Selection of Principal Component Method

Three methods were tested before the final method was selected for producing the RAI. First, a RAI was produced from the first principal component, which by definition accounted for the greatest variance in the variables. Next, the components were rotated using a varimax procedure; the significant components were identified; and a RAI was produced from the component that most clearly represented the concept of the RAI. The third method combined the significant principal components and then weighted each component according to its relative contribution to produce a combined RAI. The results of these three methods are presented in Table 4.2.

The first principal component (Table 4.2, Column 1) and rotated component (Table 4.2, Column 3) RAIs were compared to each other and to the R&D expenditures ranking throughout this paper. Both methods that combined components were discarded. A discussion of the component that was finally selected is presented in the validity section of this paper.

Comparison of the combined component (unrotated) RAI (Table 4.2, Column 2) to the first principal component RAI (Table 4.2, Column 1) indicated that an unreasonable ranking was produced. For example, the combined RAI for Stanford was 17 while the first principal component RAI was a more reliable and apparently valid (face validity) 9. Likewise, the University of California-Berkeley was ranked 18 versus 8; the University of California-Los Angeles was ranked 12 versus 6; and Indiana University was ranked 4 versus 14. In these few examples the first principal component RAI produced a more reliable and apparently valid, or "realistic" ranking. Therefore, the combined component solution was discarded.

Finally, examination of the rotated combined RAI (Table 4.2, column 4) reveals that it was almost identical to the first principal component RAI (Table 4.2, column 1). This was understandable in that by definition, the first principal component accounts for as much of the total variance as possible (Daultrey 1976, p. 3). The rotation was, in effect, taking the first principal component and the remaining weaker components and redefining them into new components. When the rotated components were recombined, they were almost equal to the original first principal component. Since little new information was revealed by the combined rotated components, the method was dropped.

Table 4.2. Comparison of the 1st PC Ranking to the Rankings of Combined Components *

University	1st. PC	PC combined	Rotd 1st. Comp.	Rotd Combined
U. of Wisconsin	1	1	4	1
Harvard	2	5	2	2
Johns Hopkins	3	6	57	3
MIT	4	8	7	5
U. of Washington	5	2	11	4
U. of Cal.-LA	6	12	3	6
Columbia	7	9	8	7
U. of Cal.-Berkeley	8	18	1	8
Stanford	9	17	6	9
U. of Florida	10	7	21	11
U. of Texas	11	10	10	12
U. of So. Cal.	12	3	31	10
U. of Illinois	13	14	5	13
U. of Indiana	14	4	67	14
Yale	15	16	12	15
Purdue	17	11	19	17
NYU	18	13	17	18
U. of Iowa	23	15	33	23
U. of Penna.	16	24	9	16
U. of Michigan	19	20	13	19
U. of So. Carolina	27	41	14	27
U. of Cal.-Davis	25	34	15	25

* The entire list of universities are not reproduced here. These rankings were produced from a preliminary set of variables, not the final set.

In summary, a number of variables were eliminated due to low multiple correlations and component loadings. Other variables were eliminated because they did not measure research activity directly. The financial variables that were subsets of total variables were eliminated because they distorted the component scores. The mix of the variables and their respective component weightings indicated there may be some overrepresentation of student related variables, which could bias the final RAI. The method of combining the significant principal components before generating a RAI was dropped because it either produced an unrealistic RAI or provided no new information.

RAI Scores

The final first principal component score rankings for all 68 universities are presented in Table 4.3. The RAI is based upon the final eleven research variables shown in Table 4.1. The rotated first component score rankings are presented in Table 4.4. Both RAIs and the R&D expenditure rankings, the traditional method for ranking research universities, are compared in Table 4.5. The next section discusses these results and in the process examines the validity and reliability of the RAIs.

Table 4.3. Research Activity Index Ranking-First Principal Component Method

FICE	University	Rank	RAI Score
2155	Harvard Univ.	1	3.427
3895	Univ. of Wisconsin-Madison	2	2.806
1312	Univ. of Cal.-Berkeley	3	2.568
1315	Univ. of Cal.-Los Angeles	4	2.145
2178	Mass. Inst. of Technology	5	1.575
3798	Univ. of Washington	6	1.534
1775	Univ. of Ill.-Urbana	7	1.266
1305	Stanford Univ.	8	1.250
2077	Johns Hopkins Univ.	9	1.242
1426	Yale Univ.	10	1.156
2707	Columbia Univ.-Main Division	11	1.147
1328	Univ. of Southern Cal.	12	.949
3658	Univ. of Texas at Austin	13	.946
2758	New York Univ.	14	.610
1535	Univ. of Florida	15	.569
3378	Univ of Pennsylvania	16	.559
2290	Michigan State Univ.	17	.526
1809	Indiana Univ.-Bloomington	18	.288
1317	Univ. of Cal.-San Diego	19	.287
1083	Univ. of Arizona	20	.282
1313	Univ of Cal.-Davis	21	.256
2520	Washington Univ.	22	.189
1892	Univ. of Iowa	23	.185
1774	Univ. of Chicago	24	.135
2974	Univ. of N.C. at Chappel Hill	25	.098
2520	Univ. of MD-College Park	26	-.050
1319	Univ. of Cal.-San Francisco	27	-.059
2130	Boston Univ.	28	-.100
2920	Duke Univ.	29	-.101
1739	Northwestern Univ.	30	-.168
1869	Iowa ST. U. of Sci.&Tech.	31	-.200
3675	Univ. of Utah	32	-.208
2627	Princeton Univ.	33	-.289
3800	Washington State Univ.	34	-.302
1598	Univ. of Georgia	35	-.350
2894	Univ. of Rochester	36	-.366
3210	Oregon State Univ.	37	-.368
3024	Case Western Univ.	38	-.383
1610	Univ. of Hawaii-Manoa	39	-.391
2972	N.C. State Univ. at Raleigh	40	-.403
3371	Temple Univ.	41	-.431
2516	Univ. of Missouri-Columbia	42	-.432
3754	Virginia Polytech Inst&St. U.	43	-.465
2329	Wayne State Univ.	44	-.475
2221	Univ. of Mass. at Amherst	45	-.539

Table 4.3--Continued.

FICE	University	Rank	RAI Score
1350	Colorado State Univ.	46	-.662
1314	Univ. of Cal.-Irvine	47	-.675
1536	Univ. of Miami	48	-.687
3242	Carnegie Mellon Univ.	49	-.690
2565	Univ. of Nebraska-Lincoln	50	-.701
3535	Vanderbilt Univ.	51	-.741
1444	George Washington Univ.	52	-.759
1052	Univ. of Alabama-Birmingham	53	-.764
1928	Kansas St. U.-Ag&Appl. Sci.	54	-.776
1131	California Inst. of Tech.	55	-.784
2903	Yeshiva Univ.	56	-.799
1316	Univ. of Cal.-Riverside	57	-.809
8789	Syracuse Univ.	58	-.810
3827	West Virginia Univ.	59	-.828
1489	Florida State Univ.	60	-.834
1445	Georgetown Univ.	61	-.913
1431	Univ. of Deleware	62	-.988
3735	Virginia Commonwealth Univ.	63	-1.008
3677	Utah State Univ.	64	-1.041
3401	Brown Univ.	65	-1.078
2423	Mississippi State Univ.	66	-1.081
2807	Rockefeller Univ.	67	-1.137
3127	Univ. of Dayton	68	-1.349

Table 4.4. Research Activity Index Ranking-Rotated First Component Method

FICE	University	Rank
2155	Harvard Univ.	1
1312	Univ. of Cal.-Berkeley	2
1315	Univ. of Cal.-Los Angeles	3
3895	Univ. of Wisconsin-Madison	4
2178	Mass. Inst. of Technology	5
1305	Stanford Univ.	6
1426	Yale Univ.	7
1775	Univ. of Ill.-Urbana	8
2707	Columbia Univ-Main Division	9
3798	Univ. of Washington	10
3658	Univ. of Texas at Austin	11
3378	Univ. of Pennsylvania	12
2785	New York Univ.	13
1317	Univ. of Cal-San Diego	14
1313	Univ. of Cal-Davis	15
2290	Michigan State Univ.	16
2974	Univ. of NC at Chapel Hill	17
1774	Univ. of Chicago	18
2130	Boston Univ.	19
2520	Washington Univ.	20
1535	Univ. of Florida	21
2920	Duke Univ.	22
1319	Univ. of Cal-San Francisco	23
1083	Univ. of Arizona	24
1892	Univ. of Iowa	25
1328	Univ. of Southern Cal.	26
1739	Northwestern Univ.	27
2103	Univ. of Maryland-Coll. Park	28
2627	Princeton Univ.	29
3371	Temple Univ.	30
3675	Univ. of Utah	31
2221	Univ. of Mass. at Amherst	32
2516	Univ. of Missouri-Columbia	33
1610	Univ. of Hawaii-Manoa	34
2894	Univ. of Rochester	35
1598	Univ. of Georgia	36
1869	Iowa St. U of Sci&Tech.	37
3024	Case Western Reserve Univ.	38
3210	Oregon State Univ.	39
3800	Washington State Univ.	40
1536	Univ. of Miami	41
3754	Virginia Polytech Inst&St. U.	42
2329	Wayne State Univ.	43
1809	Indiana Univ.-Bloomington	44
2972	N.C. State Univ. at Raleigh	45

Table 4.4--Continued.

FICE	University	Rank
2565	Univ of Nebraska-Lincoln	46
1489	Florida State Univ.	47
3535	Vanderbilt Univ.	48
1350	Colorado State Univ.	49
1314	Univ. of Cal-Irvine	50
2077	Johns Hopkins Univ.	51
2903	Yeshiva Univ.	52
1444	George Washington Univ.	53
3242	Carnegie-Mellon Univ.	54
1928	Kansas St. U-AG. & App. Sci.	55
1445	Georgetown Univ.	56
1052	Univ. Alabama-Birmingham	57
3827	West Virginia Univ.	58
1316	Univ. of Cal-Riverside	59
2807	Rockefeller Univ.	60
1131	California Inst. of Tech.	61
3401	Brown Univ.	62
3677	Utah State Univ.	63
3735	Virginia Commonwealth Univ.	64
2423	Mississippi State Univ.	65
8789	Syracuse Univ.	66
1431	Univ. of Deleware	67
3127	Univ. of Dayton	68

Table 4.5. Comparison of R&D Expenditures, RAI Rotated, and RAI First Principal Component Rankings

FICE	University	R&D$	RAI Rot'd	RAI 1st. PC
2077	Johns Hopkins	1	51	9
2178	Mass. Inst of Technology	2	5	5
3895	Univ. of Wisconsin-Madison	3	4	2
1317	Univ. of Cal.-San Diego	4	14	19
1305	Stanford Univ.	5	6	8
3798	Univ. of Washington	6	10	6
2707	Columbia Univ.-Main Division	7	9	11
2155	Harvard Univ.	8	1	1
3378	Univ. of Pennsylvania	9	12	16
1312	Univ. of Cal.-Berkeley	10	2	3
1315	Univ. of Cal.-Los Angeles	11	3	4
1775	Univ. of Ill. Urbana	12	8	7
3658	Univ. of Texas at Austin	13	11	13
1328	Univ. of Southern Cal.	14	26	12
1319	Univ. of Cal.-San Francisco	15	23	27
2290	Michigan State Univ.	16	16	17
1426	Yale Univ.	17	7	10
1083	Univ. of Arizona	18	24	20
1313	Univ. of Cal.-Davis	19	15	21
2894	Univ. of Rochester	20	35	36
8732	Purdue Univ.	21	20	22
1774	Univ. of Chicago	22	18	24
2785	New York Univ.	23	13	14
1535	Univ. of Florida	24	21	15
1598	Univ. of Georgia	25	36	35
1869	Iowa St. U of Sci & Tech.	26	37	31
3210	Oregon State Univ.	27	39	37
1131	California Inst. of Tech.	28	61	55
2972	N C State Univ. at Raleigh	29	45	40
2903	Yeshiva Univ.	30	52	56
1739	Northwestern Univ.	31	27	30
1536	Univ. of Miami	32	41	48
3024	Case Western Reserve Univ.	33	38	38
1350	Colorado State Univ.	34	49	46
1610	Univ. of Hawaii-Manoa	35	34	39
2103	Univ. of Maryland-College Pk.	36	28	26
2920	Duke Univ.	37	22	29
2974	Univ. of NC at Chapel Hill	38	17	25
1892	Univ. of Iowa	39	25	23
2807	Rockefeller Univ.	40	60	67
3675	Univ. of Utah	41	31	32
2516	Univ. of Missouri-Columbia	42	33	42
3754	Virginia Polytech Inst&St. U.	43	42	43
2565	Univ. of Nebraska-Lincoln	44	46	50

Table 4.5--Continued.

FICE	University	R&D$	RAI Rotd	RAI 1st. PC
3800	Washington State Univ.	45	40	34
1052	Univ. Alabama-Birmingham	46	57	53
1809	Indiana Univ.-Bloomington	47	44	18
3242	Carnegie-Mellon Univ.	48	54	49
2423	Mississippi State Univ.	49	65	66
2627	Princeton Univ.	50	29	33
1928	Kansas St. U-AG & App. Sci.	51	55	54
1316	Univ. of Cal-Riverside	52	59	57
3677	Utah State Univ.	53	63	64
2130	Boston Univ.	54	19	28
1314	Univ of Cal.-Irvine	55	50	47
1489	Florida State Univ.	56	47	60
2329	Wayne State Univ.	57	43	44
3535	Vanderbilt Univ.	58	48	51
3127	Univ. of Dayton	59	68	68
1444	George Washington Univ.	60	53	52
3371	Temple Univ.	61	30	41
3735	Virginia Commonwealth Univ.	62	64	63
3401	Brown Univ.	63	62	65
3827	West Virginia Univ.	64	58	59
1445	Georgetown Univ.	65	56	61
1431	Univ. of Deleware	66	67	62
2221	Univ. of Mass at Amherst	67	32	45
8789	Syracuse Univ.	68	66	58

Validity

The validity of the RAIs was established by (1) interpreting the components (content validity), (2) analyzing the impact of each variable on the rankings, and (3) analyzing the institution's rankings for reasonableness (face validity). Also, the review of the statistical methods by the expert statistician is presented.

Interpretion of the Components

The first principal component (See Table 4.6) was defined by nine variables which had coefficients that loaded heavily on the first component. Only the PCTCHG (average percentage change in R&D expenditures from 1977-1980) and the part-time students coefficients did not contribute heavily to the first principal component. The first principal component was labeled as the activity/size component. The second component was defined by PCTCHG, capital expenditures, R&D expenditures, and full-time scientists. The second component was labeled the expenditures component. The correlation matrix (See Appendix C) showed that the PCTCHG variable was related only to R&D expenditures, capital expenditures, and staff size. It appears that larger spending increases occurred at large spending institutions with large staffs. The third component was defined by part-time students and, to a lesser extent, capital expenditures, PHD awards, and full-time students. This ambiguous component was not labeled. The correlation matrix showed that the part-time students variable was related to capital expenditures and PhDs awarded and less strongly to part-time scientists and full-time students.

Table 4.6. Principal Component Analysis Score Coefficients

Variable	Name	Comp. 1	Comp. 2	Comp. 3
PCTCHG	R&DPct	.007	.565	-.175
VAR11	R&D Ex.	.134	.274	-.227
VAR193	Cap. Ex.	.097	.312	.233
INST1	FT-Sci.	.148	.217	-.085
INST3	PT-Sci.	.138	-.174	.014
SINST1	FT-Stu.	.152	-.067	.172
SINST2	PT-Stu.	.019	.122	.692
SINST3	Postdoc	.148	-.069	-.214
SINST4	Other Stu.	.146	-.220	-.230
PHD	Awards	.149	-.094	.226
ARLIND	ARL Index	.152	-.099	.078

In terms of the rotated components (See Table 4.7), component one, which was defined by eight variables, could be labeled the activity/size component. Only PCTCHG, capital expenditures and part-time graduate students did not contribute heavily to the first component. The second rotated component loaded heavily on the PCTCHG in R&D expenditures and R&D expenditures. The second rotated component loaded moderately on capital expenditures and full-time scientists and engineers. This component was labeled the expenditures component. The third and last significant component (with an eigenvalue greater than 1) loaded heavily on part-time students and moderately on capital expenditures. This component was called the eclectic component.

The first principal component and the first rotated component were very similar. The second and third components of each method were only moderately less similar. There appeared to be no particular advantage of one method over the other. The content of the components appeared to make sense.

Table 4.7. Rotated Component Analysis Score Coefficients

Variable	Name	Comp. 1	Comp. 2	Comp. 3
PCTCHG	R&DPct	-.174	.564	-.043
VAR11	R&D Ex.	.046	.350	-.141
VAR193	Cap. Ex.	-.027	.257	.307
INST1	FT-Sci.	.070	266	-.014
INST3	PT-Sci.	.188	-.119	-.006
SINST1	FT-Stu.	.156	-.057	.172
SINST2	PT-Stu.	-.060	-.058	.698
SINST3	Postdoc	.174	.038	-.202
SINST4	Other Stu.	.224	-.097	-.251
PHD	Awards	.160	-.096	.218
ARLIND	ARL Index	.172	-.063	.074

Analysis of variables impact on rankings

Next, six institutions that showed the largest differences between the RAI rankings and the R&D expenditure ranking were selected for further analysis to determine (1) which variables accounted for most of the differences between the first principal component and the rotated RAIs and (2) which variables accounted for the differences in ranking between the RAIs and the R&D expenditures ranking. From these analyses a better understanding of which variables were important in determining the final RAIs and which non-financial variables accounted for a major part of the differences between the RAI and R&D expenditure rankings, was reached.

Table 4.8 contains, for six selected institutions, each variable's mean, score coefficient, and value, for each RAI method. The selection of the six institutions was based upon a criterion that institutions must have rank order differences of 5 or more between the first principal component and the rotated component RAIs. For each institution, each variable value was compared to the corresponding mean value for all institutions. The mean value served as a reference point, as did the other institution's variable values, to help determine how a variable impacted the total component score (RAI score). Also, the component score coefficient for each variable was examined to determine whether the loading was small or large relative to the other variables component coefficients. If the component score coefficient was small, then the corresponding variable value, no matter how large, generally had relatively little impact on the final component score (RAI) for the institution.

Table 4.8. Selected Institutions for Analysis of Variable Values *

Variable Name	Mean	1st. PC	1st. Rtd Compt.	Harvard	Yale	NYU
R&D Pct	114	.007	-.174	111	113	108
R&D Ex.	53398	.134	.046	100901	71446	56739
Cap. Ex.	6099	.097	-.027	13910	9278	9044
FT-Sci.	1185	.148	.070	3674	1844	1409
PT-Sci.	241	.138	.188	1893	315	641
FT-Stu.	1632	.152	.156	2589	1570	1294
PT-Stu.	464	.019	-.060	98	34	2008
Postdoc	180	.148	.174	882	586	160
O.-Stu.	35	.146	.224	195	83	39
PHD	234	.149	.160	331	286	432
ARLIND	-.14	.152	.172	3.12	2.19	.91

	Harvard	Yale	NYU
RAI 1st. Principal component Ranks	1	10	14
RAI Rotated Ranks	1	7	13
R&D Expenditure Rankings	8	17	23

* Institutions with rank order differences of 5 or more between the first principal component and rotated component RAIs.

Table 4.8--Continued *

Variable Name	Mean	1st. PC	1st. Rtd Compt.	Johns Hopkins	U. Fla.	UC. La-Jolla
R&D Pct	114	.007	-.174	171	110	111
R&D Ex.	53398	.134	.046	253204	56656	124830
Cap. Ex.	6099	.097	-.027	11021	21808	799
FT-Sci.	1185	.148	.070	3138	2063	948
PT-Sci.	241	.138	.188	1653	153	140
FT-Stu.	1632	.152	.156	1336	2235	1043
PT-Stu.	464	.019	-.060	102	299	57
Postdoc	180	.148	.174	308	119	377
O.-Stu.	35	.146	.224	535	20	93
PHD	234	.149	.160	180	265	172
ARLIND	-.14	.152	.172	-.41	.80	.12

	Johns Hopkins	U. Fla.	UC. La-Jolla
RAI 1st. Principal Component Ranks	9	15	19
RAI Rotated Ranks	51	21	14
R&D Expenditure Rankings	1	24	4

* Institutions with rank order differences of 5 or more between the first principal component and rotated component RAIs.

The first institution analyzed was Harvard which ranked first on both of the RAI indices and eighth on the R&D expenditures ranking. In terms of the RAIs, Harvard showed superior strength on the personnel, student (except part-time) and ARL variables, all of which seemed to justifiably put it in first place. This example demonstrated a limitation of the R&D expenditures ranking.

Yale's RAI rankings were ten for the first principal component and seven for the first rotated component RAI. This differential appeared to be caused by some of the heavier non-financial component loadings of the rotated component, combined with relatively large data values (as compared to the means) for the part-time scientists, post-doctorate, other student, PHD and ARL variables. The differences between the RAI rankings and the R&D expenditures ranking of seventeen were the result of the relatively low R&D expenditure values on one hand, and strong non-financial variable values on the other. The three ranking schemes behaved as expected in that with relatively weak expenditure values, the first principal component RAI was ranked higher than the rotated RAI, which was more dependent on non-financial variables. In comparing Yale's variable values to those of like ranked institutions, it was clear that the RAI rankings presented a more reasonable ranking for Yale than did the R&D expenditures ranking. Which of the two RAIs was more appropriate for Yale was not as easily decided.

New York University (NYU) had RAIs of fourteen and thirteen, respectively. The closeness of the two may have been due to the R&D expenditures variable being close to the mean and therefore not exerting much influence on either RAI. The R&D expenditures rank was twenty-three, reflecting the relatively low average level of expenditures. The higher RAIs appeared justified because many of NYU's non-financial variables greatly exceeded the mean values. Also, the overall profile of NYU's variable values more closely represented those of institutions in the low teens of the RAI range than those in the twenty's range.

Johns Hopkin's had RAIs of nine and fifty-one. Obviously, such a difference is a matter of concern. The reason for this difference appeared to be related to differences in the PCTCHG and several other variables. The PCTCHG variable for Johns Hopkins was large, (.171) indicating a rapid growth in R&D expenditures in recent years. However, Johns Hopkins rotated component coefficient was a large negative number (-.174). When the .171 is multiplied by the -.174 and summed with the other variable scores to arrive at the RAI score, it can be seen that the large variable value resulted in lowering the RAI score for the institution. Likewise, high capital expenditures greatly reduced the rotated RAI score. These findings brought into question the value of the PCTCHG variable: it did not contribute significantly to the first principal component RAI and apparently caused a distortion in the rotated RAI. The entire rotated RAI was also brought into question because several variables contributed inappropriately to the rotated rank of fifty-one for Johns Hopkins.

The R&D expenditures rank for Johns Hopkins was one. Upon comparison of the Johns Hopkins variables to those of other institutions, it was clear that the rank of one was not appropriate in terms of a more broadly based view of research activity. For example, the coefficients for Johns Hopkins were lower than those for Harvard on eight of the eleven variables; two of these were financial variables.

The University of Florida RAIs were fifteen and twenty-one, respectively. The difference in the RAI ranks were primarily due to the small emphasis on financial variables and the very large capital expenditure values which negatively impacted the rotated RAI. The capital expenditures and full-time faculty variables for the University of Florida were much higher than the mean while the corresponding component weights were low. The low component weights reduced the influence of the high variable values on the RAI score. Comparison of the variable values to those of other institutions supported the first principal component RAI of fifteen over the rotated RAI of twenty. The R&D expenditure rank was twenty-four which reflects the average expenditure level of the university. Six of the eleven variable values were above the means for all institutions, indicating that a ranking near the average for all institutions is not appropriate. The higher RAI rank more appropriately reflects where the University of Florida's ranking should be.

The University of California at LaJolla (San Diego) had RAIs of nineteen and fourteen, respectively. This was unexpected because the first principal component RAI had a greater financial content. The explanation was USD's relatively weak values for the non-financial variables. The rotated RAI reflected the unusually high post-doctorate and other scientists variable which, given the high component coefficients, yielded a high rotated RAI. Thus, the first principal component RAI provided a better fit when compared to other institutions than did the rotated RAI, which appeared to give too much weighting to the post-doctorate and other non-faculty variables. The R&D expenditures rank of four showed the limitation of the traditional measure. LaJolla was relatively weak in the personnel, student, and some of the other variable values, as compared to similar institutions scoring in the same range.

Analysis of Institutional Rankings

This analysis focused on comparing the variable values of differing institutions to determine whether the rankings appeared reasonable. Four universities were selected for analysis based upon their relatively large differences in RAI rankings (See Table 4.9).

Boston University's first principal component RAI was twenty-eight, and it's first rotated component RAI was nineteen. Duke University's RAIs were twenty-nine and twenty-two respectively. The first principal component RAIs for the two universities were almost identical. The rotated RAIs differed by three ranks, with Boston University having the lowest rank. Was this latter difference in ranks reasonable? Boston University had less in R&D expenditures, had fewer full-time scientists, had considerably more full and part-time graduate students, fewer postdoctorate students, more other scientists, more PHD graduating students and a much smaller ARL index. In all, the values for Boston U. were larger than for Duke on six out of eleven variables. Duke was penalized for relatively low component coeficients where its values were high: capital expenditures, full-time faculty and R&D expenditures. Close rankings such as these were difficult to evaluate.

Table 4.9. Selected Institutions With Large RAI Ranking Differences *

Variable Name	Mean	1st. PC	Rotd. Comp.	Boston U.	Duke	USC	Ind-iana
R&D Pct	114	.007	-.174	110	108	124	139
R&D Ex.	53398	.134	.046	24393	39066	74304	30307
Cap. Ex.	6099	.097	-.027	1573	2672	18704	16994
FT-Sci.	1185	.148	.070	917	1446	1274	1787
PT-Sci.	241	.138	.188	332	143	332	999
FT-Stu.	1632	.152	.156	1644	899	2946	1476
PT-Stu.	464	.019	-.060	792	60	4474	453
Postdoc	180	.148	.174	122	272	317	101
O.-Stu.	35	.146	.224	57	42	9	4
PHD	234	.149	.160	319	179	424	313
ARLIND	-.14	.152	.172	-.24	.28	.38	.90
RAI 1st. PC Ranks				28	29	12	18
RAI Rotated Ranks				19	22	26	44
1st. PC Scores				-.10	-.10	.95	.29
Rotated Component Scores				.18	.14	-.04	-.51

999=Missing Data.

* Institutions were selected with rank order differences of five or more between the first principal component the first rotated component RAIs.

The next comparison was between Duke and Indiana University. Duke's RAIs were twenty-nine and twenty-two, respectively. Indiana's RAIs were eighteen and fourty-four, respectively. A comparison of the variable values indicated that Duke's were smaller than Indiana's on eight out of the eleven variables. Frequently, Duke's variables were smaller by large margins. The larger first principal component RAI for Duke appeared reasonable.

For the first rotated component RAI, the rank of fourty-four for Indiana appeared to be unreasonable. Indiana's larger PCTCHG values, larger capital expenditures combined with smaller postdoctorate and other scientists values, created a radical shift to a RAI of fourty-four. Also, missing data for the part-time scientists helped to increase the rotated RAI rank. This pointed to the rotated RAI as placing too much weight on a few variables and being overly sensitive to missing data. The first principal component RAI was less sensitive to the same missing data and was more reflective of Indiana's research activity.

The University of Southern California (USC) in Table 4.9 was compared to New York University (NYU) in Table 4.6. USC's rankings were twelve and twenty-six, respectively. NYU's were fourteen and thirteen respectively. USC's variable values were larger, frequently by significant amounts, in six out of eleven variables. Three of these variables were financial variables. This accounted for USC being lower in the first principal component RAI, which was more sensitive to the financial variables. The larger differences in the rotated RAI were due to USC's having strength in variables that had weaker component scores for the rotated component. These variables were PCTCHG, R&D expenditures, capital expenditures, and part-time students. The shift in rank from twelve to twenty-six demonstrated a weakness in the rotated RAI. Clearly, USC was not comparable, for example, to Duke, which was ranked twenty-two.

Consultants Review

Dr. Chester W. Harris, a widely respected and well published expert in factor and principal component analysis, reviewed a preliminary draft of these findings. His major suggestions are presented below and have been integrated into this study.

The previous section, which interpreted the components, corresponds very closely to Dr. Harris' interpretations. According to Dr. Harris, both the first principal component and the rotated methods produced reasonably similar solutions in terms of the major components. He also noted that if the objective is to explain as much variance in the data as possible, then the first principal component is the one to use. But, if the objective is to produce new components that load only on a few variables, then rotation may be the solution to use. The components broke down into activity/size, financial and an ambiguous component.

Since an objective of this study was to avoid the unidimensionality of the R&D expenditures measure, the comprehensive activity/size component offered a satisfactory solution. Also, it was important to maximize the variance in the data accounted for in order to enhance the validity and reliability of the resulting rankings. Dr. Harris did not review the detailed analysis of the rankings, choosing instead to comment on the interpretation of the components.

Reliability

The best test of reliability would have involved developing multi-year RAIs--a task beyond the scope of this study. Instead, the influences of the variables on the reliability of the RAI were examined for a single year.

There are two sources of data used in this study: the NSF Survey data and the ARL index data. The NSF Surveys have been conducted for over a decade. NSF variable definitions have been relatively stable, with few changes over the years. The particular variables selected for this study are some of the more stable variables. For example, headcount scientists rather than FTE scientists was used because there was some controversy over the FTE definitions and the quality of the data. Likewise, the ARL index has been stable and has been published for over five years. All variables have been used in a number of research studies. By using such well known and stable variables, it may be argued that variable reliability is established and consequently that the reliability of the RAI, which is generated from the variables, is enhanced.

The reliability of the RAI was, of course, dependent upon the consistency of the variable definitions over time. Also, the RAI was sensitive to the number of institutions included in the study. The relative position of an institution could change because of changes in the mean variable values, caused by the addition or deletion of institutions over time. Consistency of using the same variables over time is important. Table 4.10 shows the sensitivity of institutional RAI rankings from adding just one variable. When the educational & general expenditures variable was added to the final eleven variables of the RAI, the rankings for three of the first ten institutions were changed although the six highest ranked universities were stable.

Table 4.10. Comparison of RAI Based Rankings Without and With the E&G Expenditure Variable

Institution Name	Rank Without E&G Exp.	Rank With E&G Exp.
Harvard	1	1
U. of Wisc.	2	2
U. of Cal.-Berkeley	3	3
U. of Cal.-LA	4	4
MIT	5	5
U. of Washington	6	6
U. of Illinois	7	9
Stanford	8	8
Johns Hopkins	9	10
Yale	10	13

Summary

There were three significant components in both the first principal component and the rotated component solutions. The components were similar in the two solutions. In each case, the first components represented size/activity, the second represented expenditures and the third related vaguely to part-time students. The financial components were discarded because they were too unidimensional and less stable. The part-time student components were too vague to be of further interest. Also, the first principal component was more comprehensive than the rotated component and was more representative of all of the variables.

The first principal component RAI was more stable and less likely to produce rankings lacking in face validity. The superiority of the RAI over the R&D expenditures ranking was likewise demonstrated. The RAI showed higher scores for institutions that were carrying on more research activity having greater financial, faculty, student, and library committments.

The first principal component RAI appeared to be superior on most counts, with the exception that on a theoretical basis it may not be as stable if variables are added or deleted. However, Dr. Chester Harris takes exception with this theoretical contention. The analysis in this paper supports Dr. Harris. In any case the problem appears to be insignificant since the first principal component RAI produced more stable rankings than the rotated RAI. The first principal component appears to be the best choice for producing the final RAI.

CHAPTER 5

SUMMARY AND IMPLICATIONS

The primary objectives of this study were to (1) develop criteria for the selection of variables that are measures of research activity, (2) develop a composite Research Activity Index (RAI) score for each research university, (3) rank order universities, using the RAI score, and (4) evaluate the strengths and weaknesses of the RAI in terms of its validity and reliability as compared to traditional measures of university research activity. In this concluding chapter, the framework, research design and findings are summarized, and the implications of this study and suggestions for further research are discussed.

Summary of Framework

Policy analysis, planning, evaluation, and management decision making are based upon the assumption that valid and reliable information is available in a form that is readily comprehensible to decision makers. Currently, information on research activity in universities is fragmented. Frequently, a number of indicators related to research must be interpreted to arrive at judgments about research universities. The composite RAI combines a number of research indicators into a single indicator which can be used for purposes outlined in the implications section of this chapter. Also, univariant measures, like R&D expenditures, tend to be less stable from year-to-year than composite measures derived from a variety of variables. The RAI is a comprehensible indicator which can provide objective information to support decision makers in a number of activities.

From a conceptual perspective the RAI can be viewed as a summary statistic that measures the scale/activity of eleven input, process, and output variables related to research. Currently, the major but not exclusive emphasis of the RAI is in the input and process areas and to a certain extent these are proxies for outputs in the long run. Hopefully, inputs would not continue to flow to organizations if the outputs were not satisfactory.

A number of indicies have received critical attention in recent years. This study draws a clear distinction between the criticized, subjective, reputational, "quality" type indicies and those such as the RAI, which are descriptive and objective. The RAI does not attempt to measure the quality of research but instead measures research activity/scale.

Summary of Research Design

To develop the RAI, criteria for the selection of universities and research variables were developed; the initial variables were reduced by examining the strength of their multiple correlation coefficients and their loadings on the principal components. Principal component analysis was selected as the statistical technique for developing the composite RAI measure. Rotated, unrotated, and combined components were evaluated before the final selection of a component approach was made. RAI scores were computed to arrive at a score for each of the 68 research universities included in this study. Principal component analysis was used as a descriptive statistic to transform linearly a set of variables into new sets of composite variables or principal components. Each variable in the principal component had a component loading which was standardized and multiplied times that variable's standardized data values for each university. The results for each university were summed into a single score which became the RAI score for that university. The RAI scores for all universities were listed in decending order to produce the RAI rankings. The methods used to validate the RAI were to (1) examine the composition of the final variables to determine if there were any apparent biases through redundancy or over-representation of similar types of variables, and (2) perform a detailed comparison of the rankings of the RAI and R&D expenditures. This comparison was not performed because the R&D expenditures ranking was considered comparable but because it provided an understood point of reference and was useful in developing an understanding of the RAI rankings. Because of the political sensitivity of the results, a noted statistician reviewed the statistical methods as applied to the RAI.

The reliability of the RAI was partially determined by examination of the stability of the data bases and variable's over time, by examination of the variables for stability in other related studies, and by determining that like institutions generated like RAIs.

Summary of Findings

These findings relate to whether a composite RAI ranking can be developed and whether the RAI ranking is more valid and reliable than the R&D expenditures ranking. The R&D expenditures is the traditional way of ranking research universities.

Development of the RAI

It was found that a composite RAI could be developed. Out of an initial population of 112 research organizations, 68 research universities survived all editing procedures. Some of these 112 were not universities and others had either incomplete data or data that included more than one campus. Out of an initial twenty-six variables, eleven were selected for the final RAI. Variables were discarded due to low multiple correlations with other variables and/or small loadings on the principal component. Some variables were discarded because they did not directly measure research activity. Of the final eleven variables, possibly three could be discarded from future RAIs without significant, negative effects. These are the percentage change in R&D expenditures, capital expenditures, and part-time student variables.

Validity of the RAI

The validity of the RAI was determined by (1) selecting the principal component which was the most comprehensive component (content validity), (2) analyzing the RAI scores to determine the relative impact of individual variables and contrasting the RAI rankings to the traditional R&D expenditures ranking (content and face validity), and (3) analyzing the rankings for reasonableness (face validity).

The first principal component was selected for developing the final RAI. It was defined by nine of the eleven variables, all of which had similar component loadings (see Table 4.6). The other components were either dominated by one type of variable (ie. financial) or were not readily interpretable. The unidimensional financial components would most likely suffer from the same instability problems associated with the R&D expenditures variable. Rotated components and/or combined components did not produce a result that accounted for as much variance in the data as the first principal component, nor did they provide a comprehensive and balanced profile of component weights.

Detailed analysis of six universities with large variances in their first principal component RAI rankings as compared to their rotated RAI rankings and/or R&D expenditures rankings, revealed major problems with some of the rankings. By analyzing these extreme variances, major ranking problems were revealed.

Harvard showed superior strength in R&D expenditures, faculty, full-time student, postdoctorate, other student, Phd, and the ARL index variables. Harvard was clearly ranked number one. Harvard's RAI score was 3.427 and the closest other university was 2.806. However, the R&D expenditures ranking for Harvard was an unrealistic eight; this demonstrates the value of a multidimensional measure, such as the RAI.

Yale was found to be relatively stronger on non-financial variables; this accounts for its being ranked ten on the RAI, compared to only seventeen on the R&D expenditures ranking. Yale exceeded the mean variable values for all universities on eight of the eleven variables. The higher RAI score for Yale appears justified based upon the data.

New York University had an RAI rank of fourteen and a R&D expenditures rank of twenty-three. Its R&D expenditures were near the mean for all universities. Seven of its other variables were above the mean for all universities; this supported the RAI ranking over the R&D expenditures ranking as being more representative of the institution's research activity.

Johns Hopkins university was an anomaly for it had an R&D expenditure ranking of one, a first principal component RAI rank of nine and a rotated RAI rank of fifty-one! Johns Hopkins had very large financial variable values, but since the rotated RAI had small component loadings for financial variables, the rotated RAI rank was low. The same situation existed for full-time scientists. Based upon the obviously inappropriate rotated component ranking of fifty-one, the rotated method was dropped from consideration as a viable component for developing the RAI. It also was clear that when Johns Hopkins' variable values were compared to those of other institutions, the R&D rank of one was not appropriate either. This further supports the finding that a broader based measure provides a more valid indicator of research activity/size.

The University of California-LaJolla had a RAI rank of nineteen and a R&D expenditure rank of four. The rank of four was due to the R&D expenditures being more than twice the mean for all universities. However, seven of the university's eleven variable values were below their respective means for all universities, indicating that the rank of four may be too high.

In the next analysis four universities were compared to determine if universities with similar variable values were producing similar RAI rankings. If they did not produce similar rankings then the validity of the RAIs would be questioned: if the inputs are similar, the outputs (RAI) should be similar.

Boston and Duke university's variable values were compared. Boston's values were greater than Duke's on six of the eleven variables and their respective RAI ranks were an almost identical twenty-eight and twenty-nine. Likewise, Duke and Indiana universities were compared and Duke was higher than Indiana on only three of the eleven variables. As expected, Duke's RAI rank was a lower twenty-nine as compared to Indiana's eighteen. Frequently Duke's variable values were smaller, by a large margin, accounting for the eleven rank spread between the two institutions.

The University of Southern California was compared to New York University. USC's variable values were larger, frequently by significant amounts, on six of the eleven variables. USC's RAI ranking was a predictable twelve as compared to NYU's rank of fourteen.

The above findings demonstrate that the RAI rankings are predictably consistant and variations can be readily accounted for (face validity).

It should be noted that the selection of the first principal component as the best alternative for producing the RAI is further supported by Dr. Chester Harris. Dr. Harris, a noted expert in principal component analysis, reviewed a preliminary draft of this study and indicated that the first principal component should be selected if the object is to account for as much variance in the data as possible.

Reliability of the RAI

The reliability of the RAI rests upon the consistency of the RAI as shown in the previous section and the known reliability of the input variables. Numerous studies were cited which used the same variables used in this study. The NSF survey data is recognized as some of the most stable and reliable national higher education data available. The NSF Survey data have been collected for over a decade. The ARL variable has been published for over five years and has gained increasing acceptability. In addition, of the variables considered, those selected are among those that were the most stable in terms of data definitions and have been subject to the least controversy regarding their definitions and data values.

Implications

The importance of research and development to the advancement of society is well established. The federal role of supporting research has been established within the last century. That federal role is largely implemented through the major research universities in the United States. Research has become central to university aspirations of excellance, prestige, and influence. Along with student instruction and public service, research is a primary mission of these universities. Governmental public policy and university decision making require valid indicators of the research enterprise. The development of indicators that can track, over time, the inputs, processes, and outputs of research is crucial to effective decision making. The Research Activity Index developed in this study is one such indicator. It combines a number of input-related, process-related, and perhaps output-related variables into one comprehensive indicator for each research university. The input indicators are variables, such as, finances, numbers of faculty, capital expenditures; the process and output-related variables are the ARL index, graduate students, and Phds awarded.

The major beneficiaries of a RAI could be (1) federal organizations such as National Science Foundation and the National Center for Educational Statistics; (2) higher education organizations like the National Center for

Higher Education Management Systems, National Association of College and University Business Officers (NACUBO), Education Commission of the States (ECS), American Council on Education (ACE), and accreditation associations; (3) research universities; and (4) prospective faculty members and students.

The beneficiaries could use the RAI to help support the processes of (1) national policy analyses, (2) university planning and management, (3) resource acquisition and allocation, (4) monitoring research status and progress towards goals, and (5) classification of research universities.

The forms of analysis could include input into models of the research enterprise and changes and trends in rankings. A RAI-FS could be developed to rank order universities within fields of science. Also, changes and trends in rank ordering of universities within fields of science could be produced. It may be possible that methods can be developed that could measure absolute changes in RAI and RAI-FS scores. These changes could be examined to determine differences in levels of research activity (quite aside from differences in rank orders).

The implications of the RAI on national policy analysis are manifold. The development of a conceptual model of the research enterprise, as previously noted, requires indicators of that enterprise. The RAI reduces a number of variables into a single measure and can be plotted over time at different levels of aggregation (ie. universities, states or regions). A RAI could also be developed for industry and other federal agencies that perform research so that all research could be incorporated into one model. Historical trends could provide the basis for projections and extrapolations into the future; these projection and extrapolations could be analyzed for their potential economic, social, and political impacts. It could be possible to simulate various "what if" questions by manipulating the RAIs or model assumptions to test various policy scenarios.

Agencies, at the national policy level, are interested in the distribution of research activity: by military versus domestic sector, private versus public sector, by geographic region, etcetera. RAIs could be developed and organizations ranked according to these categories. Trends could be examined and changes in RAI values over time could be monitored to help evaluate the effects of policy interventions. Federal policy can result in the redistribution of funds for various reasons. For example, it could be discovered using an analysis of the RAIs that region A does a far greater share of the federally funded research than region B. It could be decided politically, based upon a fairness doctrine mandate, that region B be brought up to region A's research level. A simulation could be run that would determine the mix of additional resources needed over time to have region B's RAI on parity with region A's. As resources are made available, the RAIs in region B could be monitored to determine if planned progress toward goals is being accomplished.

Another form of analysis, other than the relative ranking of universities, could be to determine over time the change in the levels of the RAI scores. By holding the number of institutions constant in the development of the RAI, changes in the level or size of the RAI scores for subsets of institutions may indicate significant increases or decreases in their levels of research activity. The same type of indicator could be developed using fields of science. When national or state policies emphasize certain fields of science, these indicators could provide important benchmarks to focus, guide and evaluate policy decisions.

National higher education organizations and associations, such as ACE, NCHEMS, NACUBO, ECS, have an interest in the development of indicators. Indicators, like the RAI, which could monitor the progress and status of research are useful for management planning and decision making. The RAI could be used for associations that need an objective criterion for membership acceptance. The ARL index is used in this manner for research libraries. The same uses could also apply to fields of science if field of science indicators are developed.

There are a number of classification systems for categorizing postsecondary education institutions. Of importance are the criteria used for classifying institutions. The RAI could be used as one objective criterion for identifying and classifying research universities or alternatively to help validate existing classification structures.

Regional accreditation associations frequently develop indicators for evaluation of universities. They also frequently require universities to perform self studies and to develop multi-year plans based upon measureable goals. The RAIs could provide an indicator that would allow comparisons among peer universities, plotting for trend analysis, and projecting and monitoring for strategic planning purposes.

The implications for university planning are possibly the most important, in that the management tools available for administration of research have been relatively limited. Research universities have a planning interest in monitoring their total, as well as, field of science research activity. Historical trends can be plotted to determine the direction a university has taken, and planning in terms of setting target RAIs can be performed. Annual RAIs can be used to monitor and evaluate progress made towards meeting planning goals. With the development of RAIs for each field of science, an important internal management tool can become available to university administrators to evaluate past trends, compare performance against peer performance, establish measureable goals, and evaluate those goals against peer and national RAIs. Analysis of RAI-FS can help identify strong or weak areas of research activity and thus help support planning, research resource acquisition, and allocation processes. It is easier to comprehend a composite indicator than a wide range of variables indicators. Of course, a composite indicator can, and should, be analyzed when unexpected occurances or detailed explanations are indicated.

Finally, there are implications for graduate students, prospective students, researchers, faculty, businesses, and others interested in having a factual basis for selecting to associate with a particular research university. If RAI rankings were published annually, accompanied by supporting data, they could provide a useful tool for evaluating a particular university's research activity. Also, if field of science RAIs were published, they would be even of greater interest because a person interested in a particular field of science would not have to rely on a university's general reputation to select a field, if other data by field were not available.

A number of areas for further research are indicated. First, multi-year RAIs could be produced and plotted for each research university. This would provide useful trend data and would help establish the reliability of the RAI. The NSF Survey definitions that allow multi-campus reporting for a given university should be modified (retroactively, if possible) to insure consistant reporting for all universities. Also, research into the desirability and feasibility of including non-science fields that are not covered by the NSF Survey should be conducted. The most important research would be to develop RAIs for each field of science because of the potential benefit to a large number of universities and other organizations. The RAIs should be published annually in a time-series format and the possibility of a RAI national type indicator, like the GNP should be seriously considered. Finally, on a larger scale, it may be possible to develop an RAI type index for international comparative analysis of research activity.

Appendix A

POPULATION OF UNIVERSITIES

The population of research universities was initially determined by rank ordering the universities by research and development expenditures using data from the National Science Foundation's 1980 survey of Scientific and Engineering Expenditures at Universities and Colleges. The first 112 institutions were selected, excluding Federally Funded Research and Development Centers (FFRDC's). These 112 were further reduced to 68 universities based upon a matching of the NSF, NCES and ARL data bases. A prior matching performed in the NCHEMS study, Financing at the Leading 100 Research Universities, was used for the match between the NSF and NCES data bases. The primary reason for elimination of universities from this study was because NCES collects data for branch campuses (in multi-campus systems) and NSF leaves it to the discretion of the universities as to whether they combine branch campuses or report each branch separately. Table A.1 shows the universities that have been eliminated from this study because there was no exact match between HEGIS and NSF data sets. Table A.2 shows the FFRDC's that were eliminated and Table A.3 shows the universities for which sufficient data were lacking.

TABLE A.1. Leading 100 Research Institutions Without an Exact Match Between HEGIS & NSF Data Sets *

Woods Hole Oceangraphic Institute
University of New Mexico
Texas A & M University
University of Texas Cancer Center
University of Virginia
Auburn University
University of Colorado
University of Connecticut
Georgia Institute of Technology
Purdue University
University of Kentucky
Louisiana State University
University of Minnesota
Rutgers, The State University of N.J.
New Mexico State University
Cornell University
Ohio State University
University of Cincinnati
Oklahoma State University
Pennsylvania State University
University of Pittsburgh
University of Michigan
SUNY at Buffalo
SUNY at Stony Brook
University of Kansas
University of Alaska Fairbanks

Note: In addition to the above, Baylor College of Medicine was dropped because it was not a university. However, if it were combined with Baylor University, as is done by most universities with medical schools, it could have been included in this study. The University of Texas Health Science Center-Dallas was dropped because it was not combined with the University of Texas. The CUNY Mount Sinai School of Medicine was dropped because it was not a university.

* Source: McCoy et al. 1981, p. 297.

Table A.2. FFRDC's Eliminated from the Study

Lawrence Livermore Laboratory
Jet Propulsion Laboratory
Los Alamos National Laboratory
Argonne National Laboratory
Brookhaven National Laboratory
Lincoln Laboratory
Lawrence Berkeley Laboratory
Plasma Physics Laboratory
Fermi National Accelerator Laboratory
Stanford Linear Accelerator Center
National Center for Atmospheric Research
Aims Laboratory

Note: The degree of FFRDC independence from research universities varies considerably. In some instances the FFRDC is virtually on a university campus.

Table A.3. Institutions Dropped Due to Incomplete Data

University of Arkansas- Fayetteville
Clemson University
University of Rhode Island- Kingston

Appendix B

QUALIFICATIONS OF DATA

1. Both NSF and NCES estimate or impute the values for variables that have not been completed on the surveys. A combination of historical and peer institution data have been used to impute data values.

2. Research activities reported in this study are those that are specifically sponsored and separately budgeted. Departmental research, which is included as part of the institutional instruction budgets, is not accounted for in this study. There is no adequate measurement of departmental research that could be applied to the research universities. Therefore, the quantity of research is underestimated by varying amounts in each university.

3. At some universities separate research foundations acted as fiscal agents for university research. This could have resulted in underestimation of research activity at those universities.

4. Research in non-science fields was not included in this study.

5. The problems of having institutions determine how they report multi-campus activities was commented on by McCoy et al., (1981, p. 39), who recommend adoption by NSF

of the campus as the reporting unit for institutions. NCES uses a reporting procedure that tends to remove any appearance of positioning game playing. This tends to improve the cross agency utility of the NSF data.

6. Table B.1 presents cases with missing data values and substituted estimations.

Table B.1. Estimated Missing Data Values

FICE	University Name	Variable Name	Value
2516	Univ. of Missouri	FT-Sci.& Engr's	1258
8789	Syracuse Univ.	E&G Expenditure	85710734
1131	Calif. Inst. of Tech.	Avg. Faculty Comp.	38100
2903	Yeshiva Univ.	Avg. Faculty Comp.	30600
1316	U. of Cal.-Riverside	Capital Exp.	13008000
1316	U. of Cal.-Riverside	All Other-Cap. Exp.	13008000
2155	Harvard	Capital Exp.	13910000
2155	Harvard	All Other-Cap. Exp.	13910000
3242	Carnegie-Mellon Univ.	Phd's Awarded	130
1319	U. of Cal.-San Fran.	ARL Index	-1.56
1131	Calif. Inst. of Tech.	ARL Index	-1.56
2972	N.C. State U.-Raleigh	ARL Index	-1.56
2903	Yeshiva Univ.	ARL Index	-1.56
2807	Rockefeller Univ.	ARL Index	-1.56
3242	Carnegie-Mellon Univ.	ARL Index	-1.56
1928	Kansas St.U.Ag.&Appl.	ARL Index	-1.56
3677	Utah State Univ.	ARL Index	-1.56
3127	Univ. of Dayton	ARL Index	-1.56
1444	George Washington U.	ARL Index	-1.56
3735	Va. Commonwealth Univ.	ARL Index	-1.56
3827	West Virgina Univ.	ARL Index	-1.56
1431	Univ. of Deleware	ARL Index	-1.56

Several techniques were used to estimate missing data values. The missing ARL data values were estimated at a value (-1.56) slightly lower than the lowest actual ARL value reported for the highest ranking institutions. Since the institutions with missing ARL values did not qualify for membership in the ARL, their estimated ARL value would be lower than any institution that did qualify. There was no reasonable way to estimate the missing ARL values. Holding the ARL index constant, tended to nullify the effect of the index on the RAI for these institutions.

Appendix C

CORRELATION COEFFICIENTS

Table C.1. Correlation Coefficients

Variable Name	Variable Number	1	2	3	4
Percentage Change	1	1.00	.39	.19	.25
R&D Expenditures	2	.39	1.00	.41	.72
Capital Expenditures	3	.19	.41	1.00	.58
Full-Time Scientists	4	.25	.72	.58	1.00
Part-Time Scientists	5	-.12	.36	.29	.59
Full-Time Students	6	-.07	.56	.40	.59
Part-Time Students	7	.04	-.01	.23	.03
Post Doctoral Students	8	-.04	.65	.36	.64
Other Students	9	-.12	.55	.17	.50
Phd Awards	10	-.09	.50	.37	.56
ARL Index	11	-.07	.48	.40	.61

Table C.1--Continued

Variable Name	Variable Number	5	6	7	8
Percentage change	1	-.12	-.06	.04	-.04
R&D expenditures	2	.36	.56	-.01	.65
Capital Expenditures	3	.29	.40	.23	.36
Full-Time Scientists	4	.59	.59	.03	.64
Part-Time Scientists	5	1.00	.53	.16	.69
Full-Time Students	6	.53	1.00	.15	.51
Part-Time Students	7	.16	.15	1.00	-.04
Post Doctoral Students	8	.69	.51	-.04	1.00
Other Students	9	.76	.63	-.13	.77
Phd Awards	10	.52	.90	.20	.50
ARL Index	11	.60	.74	.07	.66

Table C.1--Continued

Variable Name	Variable Number	9	10	11
Percentage change	1	-.12	-.08	-.07
R&D expenditures	2	.55	.50	.48
Capital Expenditures	3	.17	.37	.40
Full-Time Scientists	4	.50	.56	.61
Part-Time Scientists	5	.76	.52	.60
Full-Time Students	6	.63	.90	.74
Part-Time Students	7	-.13	.20	.07
Post Doctoral Students	8	.77	.50	.66
Other Students	9	1.00	.62	.62
Phd Awards	10	.62	1.00	.76
ARL Index	11	.62	.76	1.00

REFERENCES

American Association of University Professors. Rocky Road Through the 1980's: Annual Report on the Economic State of the Profession. Academe. Washington: AAUP, 1980.

Ashton, Arthur B. A Proposed Research Activity Index to Rank Order Major Research Universities. Unpublished proposal, Center for the Study of Higher Education, University of Arizona, Tucson, 1983.

Bowering, David J. Draft Report on the Impacts of Federal R&D Expenditures on Graduate Student Enrollments, Graduate Degree Production, and Professional Staff Size at Leading Research Universities. Science Management Corporation, Washington, D. C., August 21, 1981.

Carnegie Council on Policy Studies of Higher Education. A Classification of Institutions of Higher Education - Revised Edition. Berkeley: Carnegie Council on Policy Studies of Higher Education, 1976.

Cartter, Allan M. An Assessment of Quality in Graduate Education. American Council on Education, Washington, D. C., 1966.

Daultrey, Stu. Concepts and Techniques in Modern Geography No. 8: Principal Components Analysis. Geo Abstracts Ltd., University of East Anglia, Norwich, England, 1976.

Dolan, Patrick W. The Ranking Game: The Power of the Academic Elite. Evaluation of Higher Education Committee of the Study Commission on Undergraduate Education and the Education of Teachers. Lincoln, Nebraska: University of Nebraska Printing and Duplicating Service, 1976, p. 3.

Drew, David E. Science Development: An Evaluation Study. Technical Report No. 4, presented to the National Board on Graduate Education, Washington, D. C., June 1975.

Ellyson, Edgar J. and Jack P. Kruegar. "Predicting Federal Research Funding at Colleges and Universities". Research in Higher Education, Vol. 13, No. 2, 1980, p. 131.

Gerhan, David R. "Graduate and Professional School Ratings: Facing the Need-to-know." The Journal of Academic Librarianship, Vol. 5, No. 4, 1979.

Glower, Donald D. "A Rational Method for Ranking Engineering Programs." Engineering Education, Vol. 70, No. 8, May 1980.

Gourman, J. The Gourman Report: A Rating of Undergraduate Programs in American and International Universities. Los Angeles: National Education Standards, 1977.

Harmon, Harry H. Modern Factor Analysis. Third Edition Revised, The University of Chicago Press, Chicago, Illinois, 1976.

Harris, Chester W. "Comments on the Preliminary Draft of: A Proposed Research Activity Index to Rank Order Major Research Universities". 1983 (unpublished).

Haywood, William T. "Using Financial Indicators in the Regional Accreditation Process." Business Officer: NACUBO, February, 1980.

Jones, Lyle V., Gardner Lindzey, and Porter E. Coggeshall, editors. An Assesssment of Research Doctorate Programs in the United States: Mathamatics and Physical Sciences. National Academy Press, Washington , D. C., 1982.

Katz, David A. "The Impact of Federal Grants on Output and Employment in Universities". Subsidies to Higher Education: The Issues, Praeger Publishers, New York, N. Y., 1980.

Kim, Jae-on and Charles W. Mueller. Introduction to Factor Analysis. Sage Publications, Beverly Hills, California, 1978.

Mandel, Carol A. and Mary P. Johnson. A Compilation of Statistics from the One Hundred and Thirteen Members of the Association of Research Libraries. Association of Research Libraries, Washington, D. C., 1981, p.5.

McCoy, Marilyn, Dennis Jones and David Makowski, Visiting Scholar. Boulder, Colorado: NCHEMS, 1979 (unpublished proposal), pp. 63-75.

__________, Jack Krakower, and David Makowski. Financing of the Leading 100 Research Universities. National Center for Higher Education Management Systems, Boulder, Colorado, April 1981.

Minter, John W. and Cathleen A. Conger. "Comparing Institutional Reports of HEGIS Financial Data with AICPA-NACUBO Standards, Part III." Business Officer: NACUBO, March 1979.

Nie, H., C. Hadlai Hull, Jean G. Jenkins, Karen Steinbrenner, and Dale H. Bent. Statistical Package for the Social Sciences, second edition McGraw-Hill Book Company, New York, 1975.

Nozicka, George J. Federally Funded Research and Development at Universities and Colleges. Final Report, Vol. I and II, Moshman Associates, Washington, D. C., 1979.

NSB National Science Board. Science Indicators: 1976. National Science Foundation, U.S. Government Printing Office, Washington, D. C., 1977, p. viii.

__________. Basic Research in Mission Agencies: Agency Perspectives on the Conduct and Support of Basic Research. National Science Foundation, U.S. Government Printing Office, Washington, D. C., 1978, p. viii.

_________. Science Indicators: 1980. National Science Foundation, U.S. Government Printing Office, Washington, D. C., 1981.

NSF National Science Foundation. A Report to the President and Congress: Federal Support to Universities, Colleges, and Selected Nonprofit Institutions--Fiscal Year 1979. Survey of Science Series, 81-308, U.S. Government Printing Office, Washington, D. C., 1979.

_________. NSF Survey of Scientific and Engineering Expenditures at Universities and Colleges. Washington: Government Printing Office, 1980.

_________. NSF-NIH Survey of Graduate Science Student Support and Postdoctorals. Washington: Government Printing Office, 1980.

_________. Survey of Scientific and Engineering Personnel Employed at Universities and Colleges. Washington: Government Printing Office, 1980.

_________. Academic Science: R&D Funds: Fiscal Year 1979. NSF-81-301, Washington, D. C., Government Printing Office, 1981.

"Research and Development at Major Universities between 1975 and 1979." Tucson Arizona: Planning and Budget Office, University of Arizona, 1981 (unpublished), pp.1-75.

Roose, K.E. and C.J. Anderson. A Rating of Graduate Programs. Washington, D. C., American Council of Education, 1970.

Rummel, R. J. Applied Factor Analysis. Northwestern University Press, Evanston, Illinois, 1970, p. 112.

Schmidtlein, Frank A. and Lucie Lapovsky. "Financial Health Indicators in the State of Maryland." Business Officer: NACUBO, February 1980, p. 19.

Stubbs, Kendon. The ARL Library Index and Quantitative Relationships in the ARL. A report for the Committee on ARL Statistics, Association of Research Libraries, Washington, D. C., November 1980.

Velicer, Wayne. "The Relation Between Factor Score Estimates, Image Scores, and Principal Component Scores." Educational and Psychological Measurement, Vol. 36, Sprint 1976, pp. 149-159.

Webster, David S. "Advantages and Disadvantages of Methods of Assessing Quality." Change, vol. 13, No. 7, October 1981, p. 21.

. "America's Highest Ranked Graduate Schools, 1925-1982." Change, Vol. 15, No. 4, May/June 1983.

Went, Victor. "Measuring Financial Conditions--III: The Federal Interest in Developing Financial Indicators for Policy Purposes." Business Officer, February 1980, pp. 22-23.