DOCUMENT RESUME

ED 285 467 HE 020 603

AUTHOR Muffo, John A., Ed.; McLaughlin, Gerald W., Ed.
TITLE A Primer on Institutional Research.
INSTITUTION Association for Institutional Research.
PUB DATE 87
NOTE 141p.
AVAILABLE FROM Association for Institutional Research, 314 Stone Building, Florida State University, Tallahassee, FL 32306.
PUB TYPE Collected Works - General (020) -- Viewpoints (120)

EDRS PRICE MF01/PC06 Plus Postage.
DESCRIPTORS Academic Persistence; Budgeting; College Programs; Comparative Analysis; Dropout Research; *Economic Factors; *Higher Education; Information Needs; *Institutional Research; Needs Assessment; Peer Institutions; *Program Evaluation; School Community Relationship; Self Evaluation (Groups); *Statistical Analysis; *Student Attrition
IDENTIFIERS *Spreadsheets; Statistical Packages

ABSTRACT

The state-of-the-art in 10 areas of institutional research is considered in this collection of 10 articles. Of concern are: student attrition and retention, academic program evaluation, financial issues and economic impacts of higher education, institutional self-study and regional accreditation, peer studies of institutions, statistical packages and spreadsheets, and the institutional factbook. Titles and authors are as follows: "Principles of Longitudinal Enrollment Analysis: Conducting Retention and Student Flow Studies" (Peter T. Ewell); "Studying Student Attrition and Retention" (Patrick T. Terenzini); "Designing and Conducting Needs Assessment Studies" (Mary Katherine Wilders); "Program Evaluation in Higher Education" (Richard F. Wilson); "Budgeting and Financial Planning" (Stefan D. Bloomfield); "Economic Impact Studies" (Mary K. Kinnick, R. Dan Walleri); "Institutional Research Support of the Self Study" (Richard D. Howard, James O. Nichols, Larry W. Gracie); "Peer Institutional Studies/Institutional Comparisons" (Deborah J. Teeter, Paul T. Brinkman); "The Institutional Factbook: Key to Perception of Institutional Research and Information Dissemination on the Campus" (J. Nichols, R. Howard, Bobby H. Sharp); and "Using Statistical Packages/Spreadsheets" (Bernard D. Yancey, Maryann S. Ruddock). (SW)

for Management Research, Policy Analysis, and Planning

A Primer on Institutional Research

edited by
John A. Muffo
Gerald W. McLaughlin

Association for Institutional Research

The Association for Institutional Research
314 Stone Building, Florida State University
Tallahassee, Florida 32306
Telephone: (904) 644-4470
Jean C. Chulak, Administrative Director

CONTENTS

INTRODUCTION

What is institutional research? This question has become something of an embarrassment to those of us who have spent a number of years calling ourselves institutional researchers. Not only has the question come from more predictable sources, such as faculty and administrative colleagues, but also from less expected people such as spouses, parents, and even children. The lack of a simple answer which is easy for those outside the field to understand can be frustrating, especially when one struggles with the question year after year.

One of the better recent definitions of institutional research is that of Peterson (1985, p. 5), who calls it "a critical intermediary function that links the educational, managerial, and information functions of higher education institutions and functions." Another frequently cited description of the field is offered by Saupe (1981), who emphasizes the importance of internal data gathering and analysis to the management of institutions of postsecondary education.

An insightful summary of how the institutional research function is best defined can be found in Maassen (1986). A Dutchman studying the activity of institutional research in Western Europe, where the term itself is generally unknown in any language, he concludes that four basic activities describe institutional research:

1. Collecting data about the performance of an institution,
2. Collecting data about the environment of an institution,
3. Analyzing and interpreting the collected data, and
4. Transforming the data analyses and data interpretations into information that can be used to support institutional planning, policy making, and decision making (Maassen, 1986, p. 4).

In the U.S. as in Europe, the activity of institutional research may be conducted by a person unfamiliar with the term but very familiar with the types of activities described by Maassen. Assistant deans, registrars, and a host of others are often performing institutional researcher functions without even knowing it.

It can be seen from the preceding discussion that IR functions are processes - processes which can be, and are, performed in and on most functional areas of the institution. It is difficult, however, to consider processes without relating them to the function they most frequently perform and the product they typically produce. These functions and products are often identified as the activities of IR.

The purpose of the Primer is to provide a practical introduction to some of the more common types of activities in which institutional researchers become involved. Though the Primer was originally conceived by some as being aimed at newcomers to the field, it quickly became obvious that even the most experienced veterans are relative newcomers in certain areas of activity. Consequently we now consider the Primer as an introduction to key topical areas aimed at all who are interested. Those who require more in-depth understanding of a subject are provided with bibliographical references which can take them beyond the basic concepts and issues provided here.

We were quite fortunate in assembling 15 highly respected authors to describe the state-of-the-art in 10 areas of institutional research. Peter Ewell and Pat Terenzini lead off with chapters on student attrition and retention, examining different facets of this important topic. Kathy Wilders provides an overview of another aspect of student enrollment planning and analysis, needs assessment studies. Together the first three chapters help explain how institutional researchers can become involved in attracting and retaining students, an activity central to the health of any college or university.

The next five chapters shift the focus from the student level to the institutional or organizational level. Dick Wilson shows how various external and internal forces shape efforts to evaluate academic programs. Stefan Bloomfield sheds light on the sometimes seemingly impenetrable subject of budgeting and financial planning, an area with increasing need for sound data, analysis, and interpretation. The economic impact of an institution of higher education on the broader community is the subject of the next chapter by Mary Kinnick and Dan Wallerí; the extent of the impact in economic terms is often a great surprise to those unfamiliar with economic impact studies. Rich Howard, Jim Nichols, and Larry Gracie follow with an overview of how institutional research supports the self-study process of regional accreditations, providing a number of helpful tips in the process. Lastly among the chapters on institutional management concerns is that of Deb Teeter and Paul Brinkman regarding peer studies of institutions. Those of us whose salary bases rely in part on the results of such studies have a profound appreciation for what this chapter tells us.

The last two chapters of the Primer might be called techniques or methodology descriptions, since they provide overviews of common products (the institutional factbook) and tools (statistical packages and spreadsheets) of the trade. Jim Nichols, Rich Howard, and Bobby Sharp show how best to assemble and display a factbook, placing special emphasis on the needs of the environment in addition to the more mechanical aspects. Bernie Yancey and Maryann Ruddock conclude the volume with some helpful directions for exploring the maze of computer software likely to confront the institutional researcher, among others.

We are quite pleased with the contributions of these authors to the profession; all of us owe them our thanks for so generously sharing their knowledge and experience. Certainly there are other subjects that might have been included here, for institutional research has applied its processes over many broad areas of our institutions. This volume does not purport to be comprehensive. It is a beginning, however, in summarizing the activities of the profession.

We would like to acknowledge the efforts of Beth Oehring and Al Bloom of the Office of Institutional Research and Planning Analysis at Virginia Tech. They were instrumental in converting ten sets of different machine readable copy into a unified whole. A special thanks is due also to Ms. Cheryl W. Ruggiero of the Virginia Tech Writing Laboratory for her thorough copy editing of the text.

We also would like to express our appreciation to the Publications Board and Executive Committee of AIR for the concept and support for this effort.

What, then, is institutional research? We still don't have an easy, simple answer for spouses, parents, children, and other relatives. The Primer is something that can now be shared with our colleagues, however. We hope you find it useful and interesting.

John A. Muffo
Gerald W. McLaughlin

Virginia Tech
April, 1987

REFERENCES

Maassen, P. A. M. (1986). Institutional research and organizational adaptation. Paper presented at the Eighth European AIR Forum, Loughborough, England.

Peterson, M. W. (1985). Institutional research: An evolutionary perspective. In M. W. Peterson & M. Corcoran (Issue Eds.), Institutional research in transition. New Directions for Institutional Research (No. 46, pp. 5-15). San Francisco: Jossey-Bass.

Saupe, J. L. (1981). The functions of institutional research. Tallahassee, FL: Association for Institutional Research.

PRINCIPLES OF LONGITUDINAL ENROLLMENT ANALYSIS: CONDUCTING RETENTION AND STUDENT FLOW STUDIES

Peter T. Ewell
National Center for Higher Education
Management Systems (NCHEMS)

In an era of constrained financial resources, maintaining enrollment has become increasingly important for colleges and universities. Equally pressing for many institutions have been recent public concerns about increasing the quality of undergraduate education and about demonstrating the success rates of enrolled students (NIE, 1984; Ewell, 1984a). In response to these concerns, many colleges and universities are devoting greater attention to monitoring and shaping the characteristics of their student bodies. Recruitment programs are being more carefully targeted to increase the probability of successful program completion. Retention programs are being developed with much greater sensitivity about the need to develop different kinds of assistance and advisement for different kinds of students.

College and university administrators are increasingly using the term "enrollment management" to describe their coordinated efforts to respond to those challenges and to build and maintain an institutional enrollment profile of given size and shape (Zemsky & Associates, 1980; Kreutner & Godfrey, 1981; Kemerer, Baldridge & Green, 1982; Hossler, 1984; Lonabocker & Halford, 1984). Enrollment management programs attempt to coordinate all policies and activities that influence enrollment, and to monitor their effects. In its strongest construction, the responsibility for enrollment management may rest with a specially created office or position reporting to a high-level administrator. In its weaker forms, the function involves establishing a set of common goals among the various offices responsible for the total enrollment picture.

Regardless of its configuration, good research about the dynamics of enrollment is critical to the success of an institutional enrollment management program. The purpose of this chapter is to describe briefly some of the methods necessary to conduct adequate enrollment management research - particularly the structure, design and implementation of longitudinal student flow studies. Accordingly, the chapter will be organized in terms of three main sections:

1. Basic Principles and Concepts will describe the conceptual basis for longitudinal studies and the primary principles involved in determining enrollment structure.

2. Longitudinal File Construction will cover principles and techniques for constructing "cohort" tracking files that can be used to document and analyze student progress over time. Such files, built from existing

student records, constitute the primary data resource for longitudinal studies.

3. Questionnaire Data Collection will cover the use of survey questionnaires to supplement data from longitudinal files, and the administration of periodic surveys to an identified student population over time.

In each case, the emphasis is on articulating basic principles, and on developing an ongoing, integrated research program that can be maintained over time. A selected bibliography of basic references on several areas of enrollment management research is available as a supplement to this chapter by writing to NCHEMS.

BASIC PRINCIPLES AND CONCEPTS

The best way to begin development of an integrated research program to support enrollment management decisionmaking is to consider carefully what it needs to accomplish. At minimum, the following requirements are needed:

1. The program must be able to conceptually link all processes that contribute to the institution's overall enrollment profile. These include the admissions process, transfer, withdrawal, dismissal, reenrollment, and program completion. Stated another way, this requirement demands that student enrollment be modelled from the outset as the product of a specified longitudinal process - one that begins with a given student being identified as a member of a defined recruitment population, and that ends with unambiguous withdrawal or program completion.

2. The program must be able to identify and distinguish the behaviors of different kinds of students. Colleges and universities are made up of significantly different types of students who can behave in quite different ways. An effective research approach should therefore be able to determine (a) how many distinct behavioral groups of students are enrolled, (b) how large each is, and (c) how they behave under different circumstances with respect to patterns of persistence and completion.

3. The program must allow analysts to estimate the effects of proposed policy changes on total enrollment and on the distribution of enrollment. The primary reason for doing longitudinal studies is to establish a basis for building a model of enrollment dynamics over time. Once such a model is established, it must be manipulable to the extent that the enrollment consequences of different courses of action can be projected and compared.

4. The program must provide a framework for organizing and interpreting the results of past research about students, and for identifying the kinds of future

studies of student behavior that might be most fruitful. Much institutional research about students consists of discrete studies carried out for disparate and particular purposes, the results of which are never again used. A longitudinal model, once established, provides a mechanism for organizing the findings of past research so that they can shed light on enrollment behavior. At the same time, unknowns in the model may suggest needed future research.

Each of these requirements is easy to state, but difficult to fulfill in practice. Minimally, however, two research approaches are required: (1) construction and estimation of a comprehensive longitudinal model of student progress that can show how students of different kinds move into, through, and out of the institution, and (2) identification of a minimum number of distinct behavioral groups that together constitute the bulk of an institution's enrollment.

Satisfying the first research requirement involves building and estimating a mathematical model that represents student progress through the institution as a set of linked events and decisions (Fwell, 1984b). Over the years, many such models have been developed for enrollment projection or for estimating future tuition revenue (for example, Kraetsch, 1979-80; Wing, 1974). Most are explicitly termed "Markov" or "cohort survival" models. The essence of such a model is a set of "transition probabilities" that determine the distribution of a particular population among various defined states over a number of discrete time periods (Stokey & Zeckhauser, 1978). For students, these "states" can be defined in many ways - including active enrollment, nonenrollment through graduation, academic dismissal, or voluntary withdrawal - and "discrete time periods" are generally represented as lapsed terms or years.

Figure 1 presents an overview of such a model for a particular student cohort. The model contains distinct components for both admissions and student persistence, but the two are linked in order to estimate the respective or simultaneous impacts of changes in each of these areas. The logic of the model is to represent student progress as a series of discrete decision points through which each student must pass. At each decision point, a probability of successfully passing the decision point may be calculated from past trends, and total enrollment is a function of all such decision point probabilities. Furthermore, decision points are of two distinct types - those under the control of the student, and those determined by institutional action or policy. Matriculation rate and voluntary withdrawal are examples of the former, while acceptance rate and the rate of academic dismissal are examples of the latter. Together, these two types of decision points constitute a complete chain of events that operate in concert, and that determine the enrollment

status of a particular group of students at a particular point in time.

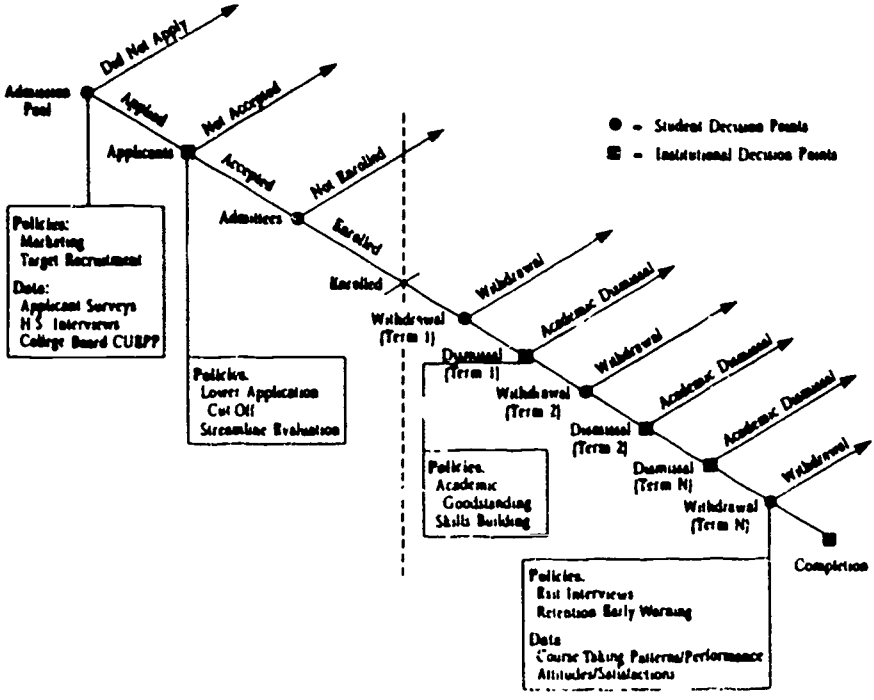


Figure 1. Conceptual Model of Student Flow Process

Such models, however, are of limited value if they do not take into account the different kinds of students typically found in an undergraduate population. Different kinds of students may behave in systematically different ways. It therefore may be necessary to build distinct tracking models (with quite different values for transition probabilities at each decision point) for different types of students. But what kinds of differences are important? How should such subpopulations be defined?

Institutional researchers traditionally break down student population in two ways - demographically and by program area. Such breakdowns are, however, generally done one at a time. Separate tracking analyses, for example, are commonly conducted for the male and female portions of a cohort, for the older and younger groups, or by department or major. While this approach will certainly provide some insight, distinct behavioral groups will more often consist of combinations of such factors. A black male who is 18 to 21 years old and seeking entry-level occupational

skills, for example, is far different from a white female liberal arts student attending part-time during the day to obtain a cultural enrichment experience.

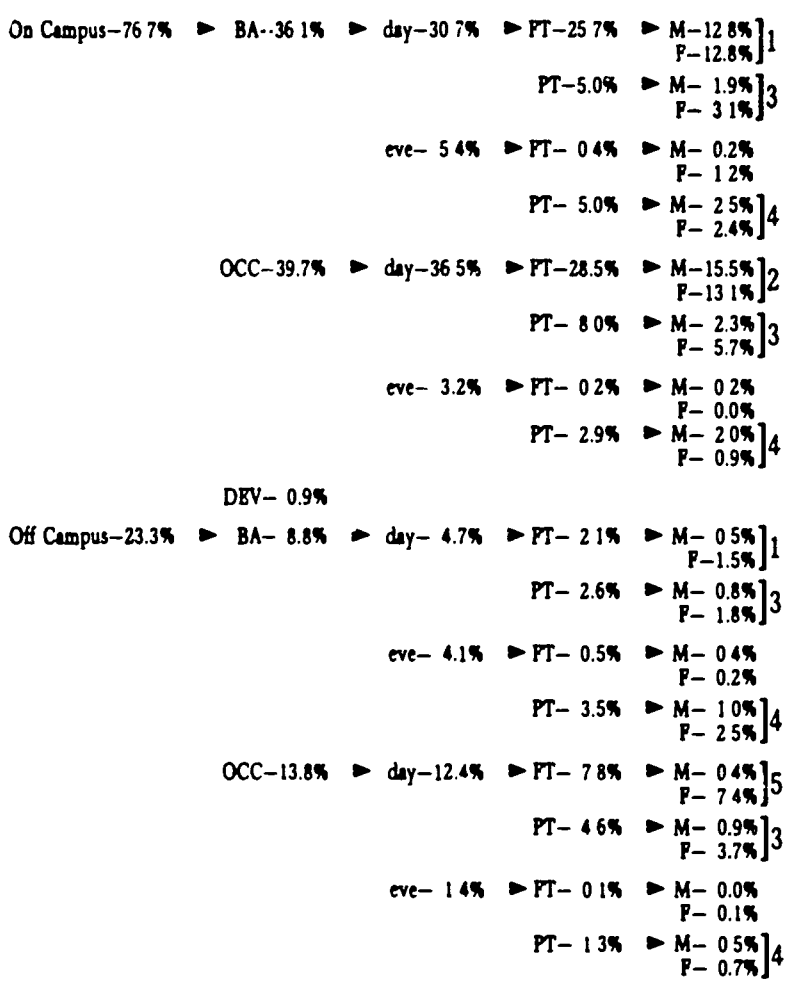
Appropriate tracking groups are therefore best identified by disaggregating total enrollment by a number of cross-cutting variables. Naturally, the choice of such variables will depend on both the nature of the institution and the characteristics of the population under study. Figure 2, for example, shows such a multiple disaggregation for a small rural community college.

The right-hand side of this breakdown represents a set of logical possibilities for cross-cuts among a set of demographic and enrollment variables. Rarely, however, will all such logical possibilities contain substantial numbers of students. Rather, students will cluster in certain categories that can then be reaggregated for analytical purposes. In the example shown, 96.2 percent of the population is accounted for by five behavioral groups. Each of these groups, once identified, was tracked separately. In order to be analytically useful, of course, behavioral groups of this kind must be statistically stable. This means that there must be sufficient numbers in each group to ensure that significant differences in enrollment behavior are detected. A good rule of thumb is to insist on group sizes of at least 100, or at least 5 percent of total enrollment. While there is no "right" way to identify such groups, experience indicates that most undergraduate populations can be usefully disaggregated into six to ten discrete tracking groups.

LONGITUDINAL FILE CONSTRUCTION

Cohort survival models are relatively easy to describe, but it may prove a challenge to obtain the data needed to estimate them. The best way to meet this challenge is to construct a set of longitudinal tracking files for particular entering "cohorts" of students - files that contain a student-by-student enrollment history for members of the cohort over a designated number of consecutive terms. The data in such files enable the analyst to answer the question "What is the enrollment pattern of each individual in the cohort?" Construction of such a file depends upon the availability of "frozen file" student record information for a number of past terms of enrollment. If "frozen files" are not immediately available, an alternative is to obtain past census date enrollment records and create them.

Basic issues involved in constructing longitudinal files are described in this section. Because of data limitations, however, analysts should be aware of two approximate methods for determining transition probabilities in a student flow model. The first uses current term information alone, and employs the student's first term of academic history at the institution to determine the number



1 = Full-time, Day, Transfer (27.1%)
2 = Full-time, Day, Occupational (28.6%)
3 = Part-time, Day, Program (20.2%)
4 = Part-time, Evening, Program (12.5%)
5 = Off-campus, Full-time, Day, Occupational (7 8%)

Figure 2. Breakdown of SRC Enrollment by Types of Students Fall, 1980

of currently enrolled "survivors" of an original entering cohort. The essence of this approach is to answer the question "How many students (of type A) entering the institution in term X were actively registered in term Y (the current term)?" While such an analysis will not yield individual student enrollment histories, it may allow approximation of the survival rates of different student populations from term to term. It is also important to note that the analysis depends upon the integrity of the first term of academic history stored in the student record; many student record systems "update" this element to reflect readmission, change in program or degree sought, etc. Changes of this kind will eliminate the possibility of using this data element to identify a student's actual first term of enrollment.

A second approximation method involves calculating a term-to-term survival estimate on the basis of total enrollment and graduation figures. Persisters from the immediately preceding term are estimated as the total number enrolled in the current term, less those graduating at the end of the immediately preceding term, less those newly admitted or readmitted at the beginning of the current term. In essence, this method answers the question "How many of those who enrolled last term also enrolled this term and did not graduate?" Both methods will allow approximate estimation of term-to-term persistence for different student population groups. Neither, however, will allow detailed investigation of the historical dynamics of enrollment flow.

The primary difficulty in building a student-by-student longitudinal tracking file is the manner in which student record data is generally stored and accessed in computerized registration files. Most student record systems maintain two quite different types of files that contain enrollment information. The first, often termed "transcript" or "student history" files, contain historical information on student demographics, course-taking, and performance. Because they are used to generate transcripts and other student records, they must contain historical information, but they must also have information on a given student's status at a particular point in time. The information in such files is thus generally inappropriate for tracking because the codes and values for such elements as probation, grade point average (GPA), and current major are constantly updated, and are written over the old values. As a result, it may be impossible to track such things as changes in student major, or to note correlations between probationary status or incomplete course work and subsequent persistence.

The alternative is to use files of the second type - so-called "term enrollment" files - but these present their own difficulties. Such files are maintained for each term, and contain records for all the students enrolled in that term - usually "frozen" as of the tenth day of the term. They are generally used for HEGIS and other summary re-

porting purposes. Usually they contain much of the same information as do history files, but the data are recorded as of that term and are never updated. This means that it is possible to capture important student status elements at a particular point in time. But it also means that updates representing error corrections - for example, corrections of incorrect or missing Social Security numbers - are not present prior to the point of correction.

The primary difficulty with term enrollment files, however, is that they are free-standing and are difficult to link together without special programming. Often, in fact, they are archived off-line and must be physically put up on the system. This means that it is sometimes a complex problem to successively read such files, to select the records for a given cohort of students, to extract the relatively few data elements from each record that are required for an analysis, and to conduct the analysis itself.

Institutions that maintain a substantial number of past term files on-line, and that also have access to a powerful statistical package such as SAS that allows simultaneous use of several source files, can accomplish these tasks in a single step. But even in this case it is often a good idea to maintain distinct longitudinal files for analytical purposes for ease of access and data manipulation, and for maintaining important derived data elements. For most institutions, therefore, the best answer will be to construct a free-standing tracking file for each cohort. This file will contain a longitudinal record consisting of extracted data elements from past term files for each member of the cohort. Figure 3 presents a typical minimal record layout for a tracking file of this kind. Creating such files first involves identifying all members of the cohort to be tracked and locating their first term record in the student database. Maintenance and updating of the file is accomplished each term by adding current term information to the end of each record. Once procedures are established, this is generally a straightforward process for a computing center. Furthermore, because the resulting file is of fixed length for all records, it is easily analyzed by SPSS or by other statistical packages that require (or that work better with) such a format.

In constructing a longitudinal file, a number of questions need to be answered.

Who will be tracked? Tracking cohorts are generally defined in terms of time of entry or first term of academic history - for example, all students who first entered the institution in the Fall of 1985. While this decision rule appears simple, a number of further decisions may need to be made. A common question is whether or not to track non-program or non-degree-seeking students. Such students are often not tracked because their enrollment periods may cover only one or two terms and examination of detailed behavior patterns is deemed unnecessary. On the other hand, non-program status occasionally precedes admission

Figure 3

STUDENT TRACKING DATABASE
PROPOSED DATA ELEMENTS

DATA ELEMENT	POSITION	LENGTH	TYPE
Social Security Number	1-9	9	N
Year of Birth	10-11	2	N
Sex	12	1	AN
Race/Ethnicity	13	1	AN
Last Prior College	14-19	6	AN
Student Intent*	20	1	AN
Entering Student Type	21	1	N
Entering Student Major	22-25	4	AN
Entering Student Time of Attendance	26	1	AN
Degree Field Attained*	27-30	4	AN

DATA ELEMENT	POSITION	LENGTH	TYPE
Term 1 Hours Attempted	31-35	5	N
Term 1 Hours Completed	36-40	5	N
Term 1 Cumulative GPA	41-43	3	N
Term 1 Degree Earned Flag	44	1	N

	TERM							
	2	3	4	5	6	7	8	9
Hours Attempted	45-49	59-63	73-77	87-91	101-105	115-119	129-133	143-147
Hours Completed	50-54	64-68	78-82	92-96	106-110	120-124	134-138	148-152
Cumulative GPA	55-57	69-71	83-85	97-99	111-113	125-127	139-141	153-155
Degree Earned Flag	58	72	86	100	114	128	142	156

*Field reserved for data currently not collected.

into a particular program if a student lacks prerequisites, has not had his or her records processed in time, etc. Similar issues surround the treatment of transfer students, and of readmitted students.

How long will they be tracked? Theoretically, a cohort can be tracked indefinitely (or at least until all members are graduated or dead!). In practice, however, it is best to establish a "cut-off" point beyond which the cohort is considered closed. Because of increasing numbers of part-time enrollees and of "stop-out" students who interrupt their enrollment, it is generally best to include sufficient time for students to finish - at least six years for a four-year program and four years for a two-year program. Indeed, the national standard for completion of an undergraduate degree is now five years. The most useful cohort tracking period will, of course, vary both by institution and by the types of students tracked. A good rule of thumb, however, is to maintain a tracking period of sufficient length to determine the fates of at least 90 percent of the students in the cohort.

How often will new cohorts be established? In principle, it is possible to begin a new tracking cohort each term. In practice, however, maintaining simultaneous tracking files for as many as twenty different cohorts can be complex. Furthermore, there is generally relatively little variation in results across cohorts from successive years, unless there has been a substantial change in the characteristics of newly admitted students or in the program or institutional environment. As a result, most institutions establish new cohorts on a periodic basis - for example, once every three years. A more complicated issue is whether to establish cohorts for terms other than fall. For institutions that admit students year-round or for graduate programs, establishing spring or winter cohorts may prove fruitful because students who first enter the institution at these times may differ significantly from fall entrants.

What data elements will be tracked each term? The actual data content of a longitudinal cohort file can vary considerably. A major question, however, is what data elements to read and record each term. At minimum, each record should contain information on enrollment status, on hours attempted and completed, and on grade/probationary information for each term. Given institutional needs, however, many other elements could be tracked. Because space must be reserved for each additional element for each term in the tracking period, however, every additional element will considerably increase the size of the tracking file.

The most complex of these choices, of course, is that of the data elements to include in the file. Four distinct types of data elements are generally typical of student tracking studies, and each is handled in a somewhat different way.

1. Fixed data elements are those which never vary. Most are demographic - for example, gender, date of birth, etc. - but they can include such elements as high school performance, last prior college, etc. Generally these elements are extracted from the term file that corresponds to the student's first term of academic history. They should be checked periodically against more recent information, however, because recent corrections are rarely posted to past term files.

2. Variable data elements are those which must be recorded and tracked each term. Examples include enrollment status, hours attempted and earned, term GPA, probationary status, and sometimes student major. These elements are read successively from historical term files throughout the tracking period, and each record will contain an entry for every term. Blanks are generally recorded in these fields for terms in which the student was not enrolled.

3. "Semi-fixed" data elements are those which occasionally vary in the course of a student's enrollment, but which do not do so systematically. These include demographic elements (for example, marital status, employment status, etc.) and enrollment elements (for example, program major, residence status, day/evening attendance, etc.). Perhaps the biggest decision involved in constructing a longitudinal tracking file is how to handle such elements. If changes in these elements are important, they may be treated as "variable," and recorded each term. Because this can vastly increase file size if many such elements are tracked, the decision is often made to treat some elements of this kind as fixed. This is often done for employment status, marital status, and other "semi-fixed" demographic elements. In this case, the value of the element is taken for the student's first term of academic history only, and is assumed to remain the same throughout the tracking period.

4. Derived elements are those that are calculated from others in the file for particular analytical purposes. Examples range in complexity from student age (derived from date of birth), to such factors as course completion rates (derived from cumulated hours attempted and earned). Derived elements are generally not physically maintained in longitudinal record files, but are calculated by statistical packages on an as-needed basis. If calculations are complex and large numbers of people are using the file for analytical purposes, it may be efficient to calculate such elements a single time and maintain them in the file.

A list of the most common data elements used in free-standing tracking files, classified by type, is presented in Figure 4.

Figure 4

Data Elements Most Commonly Used in Student Tracking Systems

Fixed/Semi-fixed:

SSN	Financial aid status
Date of birth	Entering student type
Gender	Major/program*
Race/ethnicity	Time of attendance*
Last prior college/hs	[Student intent]
Test scores	[Employment status]
	[Marital status]

Variable (Term):

Hours attempted	Academic status
Hours completed	Major/program*
Term gpa	Time of attendance*
Degree earned flag	[Requirement status]

Derived:

Total hours attempted	Still enrolled (after N terms)
Total hours completed	First term only
Average load	Completed
Completion rate	Stopped out

Once established, tracking files of this kind can be used to support many kinds of analyses. The most straightforward are simple cohort survival studies that involve calculating completion and dropout rates for different student subpopulations. More complex and revealing analyses can include multivariate predictive studies of student persistence, dropout, and performance. Such studies are generally regression-based, but may also fruitfully employ such techniques as discriminate analysis or probit analysis. A major utility of such files, of course, is to estimate a generalized Markov or cohort survival model for purposes of predicting enrollment, as described in the previous section.

USE OF SURVEY QUESTIONNAIRES IN LONGITUDINAL STUDIES

Student tracking data drawn directly from registration record files provide a powerful tool for documenting patterns of student enrollment behavior. By analyzing such patterns by subpopulation, and by correlating them with student demographics and with such factors as academic performance, time of attendance, etc., a great deal of insight can be gained about how students behave as they do. As Pat Terenzini emphasizes in his chapter, however, such inferences are necessarily limited because they do not tell us the "why" of student behavior. Administering survey questionnaires in conjunction with longitudinal tracking studies can help illuminate some of the reasons behind revealed behavior.

Such surveys are generally undertaken for one of three purposes:

1. To regularly collect information that is not currently included in registration records. Examples include employment status, marital status, and academic goal.

2. To study a particular student subpopulation in greater detail. If tracking studies reveal a particular group of students to be unusually prone withdrawal, for example, a survey might be administered to such students to determine their reasons for withdrawal and their current situation.

3. To study a particular "decision point" in a Markov or cohort survival model. If tracking studies reveal that the end of the first term is a crucial time for persistence, a survey might be administered to all students passing through this point to determine their intentions, perceptions, and attitudes.

It is important to note that the latter two uses of surveys represent considerable gains in efficiency over administering of periodic surveys to all students. Rather, surveys are used to investigate more carefully focused research questions suggested by prior behavioral studies.

As in constructing a longitudinal tracking file, survey design must consider (1) who is to be surveyed (for example, minority students, program students, evening students, etc., and whether or not samples should be employed), (2) when surveys should be administered (for example, at entry, after graduation, within a year of withdrawal, etc.), (3) how often to survey (for example, every third year, by program on a rotating schedule, etc.), and (4) what questions to ask. Typical data collection points are presented in Figure 5.

Data collected by survey must also be stored and maintained. Decisions must therefore be made on how and whether to physically integrate such information with longitudinal student data files. Some institutions reserve data fields in their longitudinal cohort files for the inclusion of responses to important survey questions. Others maintain survey information in separate files, and use a linking element (such as student number) to tie survey responses to enrollment behavior. Still others maintain no link between such files, and simply use questionnaire data to explore hypotheses suggested by longitudinal tracking studies.

As in the establishment of record-based longitudinal files, the primary question in using surveys is what to include. Typical questions asked on tracking surveys depend upon the point in a student's academic career at which information is desired. Questions typically asked at entry include (1) student goal and intended duration or persistence, (2) information about college choices including reasons for choosing this college and sources of information important in making the decision to enroll, and (3) perceived readiness and need for remediation. Typical questions asked of currently enrolled students include (1) progress toward fulfilling goals, (2) knowledge of, use, and evaluation of particular programs and services, (3) perceptions of and involvement with the campus environment, (4) perceived gains in knowledge and skills, and (5) perceived changes in attitudes and beliefs. Questions typically asked of former students (both successful graduates and non-completers) include (1) goal fulfillment, (2) subsequent attendance at other colleges and universities, (3) subsequent and current employment situation, (4) knowledge of, use, and evaluation of particular programs and services, (5) perceived strengths and weaknesses of instruction given current circumstances, and (6) reasons for persistence or withdrawal.

Questionnaire studies of this kind are often facilitated by using a pre-designed instrument, by using "item-banks" containing many preconstructed questionnaire items for use in local surveys, and by reviewing the literature on student choice and persistence. An advantage of using commercial surveys is that comparative information from other, similar institutions is often available. Most such instruments also include space for up to twenty items designed locally. Figure 6 provides a list of the most

FIGURE 5

STUDENT-OUTCOMES DATA COLLECTION POINTS FOR COLLEGES AND UNIVERSITIES

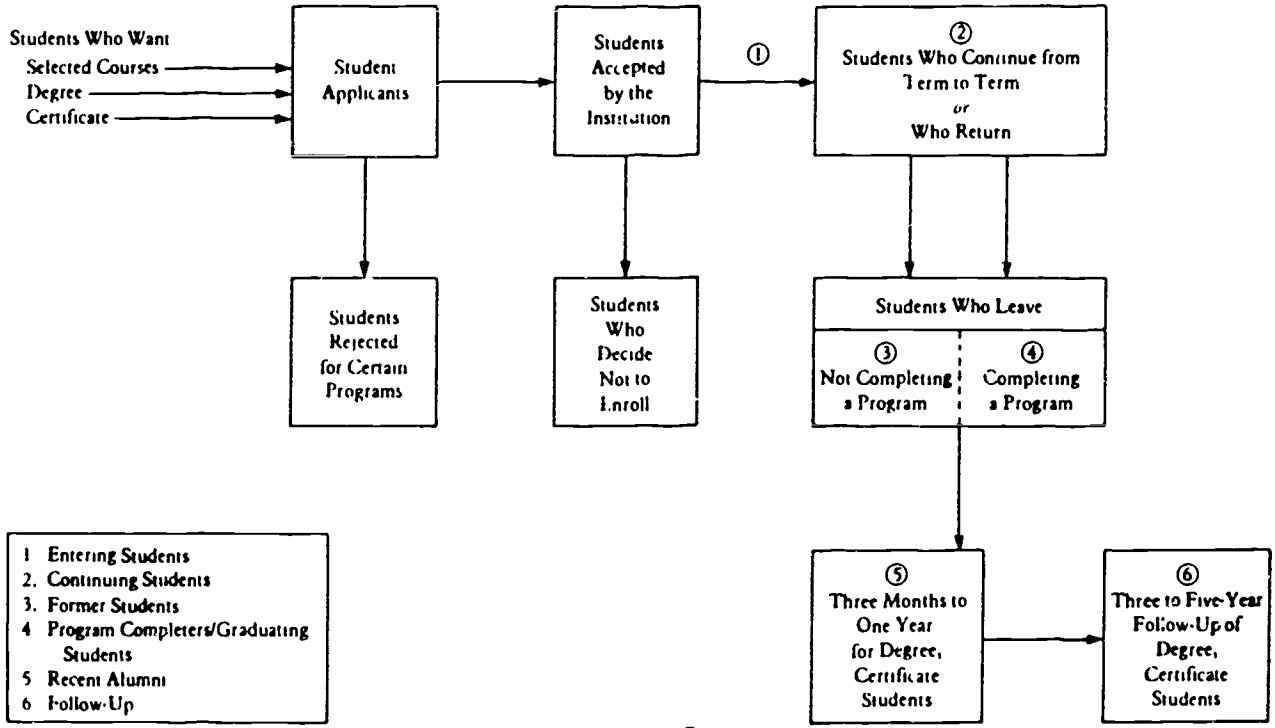


Figure 6

Sources of Questionnaires and Questionnaire Items

Existing Major Student Questionnaire Systems:

• Cooperative Institutional Research Program (CIRP)	Higher Education Research Institute - UCLA
• ACT Evaluation Survey Service (ESS)	ACT; Iowa City, Iowa
• Student Outcomes Information Service (SOIS)	NCHEMS; Boulder, Colorado
• College Student Experiences Questionnaire (CSEQ)	Higher Education Research Institute - UCLA

Sources of Items/Ideas for Local Questionnaires:

• "Surveying Your Alumni" (McKenna 1983)	CASE, Washington, DC
• FIPSE Project Item Bank	Chancellor's Office, California Community Colleges
• Higher Education Measurement and Evaluation Kit (Pace 1975)	Higher Education Research Institute - UCLA
• Outcome Measures and Procedures Manual	NCHEMS; Boulder, Colorado

common commercial student surveys and some useful sources for items when designing local questionnaires.

It is important to emphasize that questionnaire surveys pay their greatest dividends when designed in concert with record-based student tracking studies. Because of problems in obtaining adequate response rates and because of the inherent ambiguity of all survey data, it is best to begin analyzing enrollment behavior by making maximum use of existing student record information. Survey data can then be used to help explain or to illuminate the patterns of enrollment obtained for particular types of students. Survey data alone, however, can be notoriously unreliable in forecasting or accounting for enrollment behavior. Survey data may, for example, reveal that students are particularly unhappy with academic advisement, with parking availability, or with a range of other aspects of the campus environment. But the frequency of such perceptions may be as great for persisters as for non-persisters. Documenting actual behavior must precede interpreting perceptions of this kind, as the interpretation will vary considerably depending upon what students actually ended up doing as a result.

CONCLUSION

In conclusion, it is important to re-emphasize the fact that an ongoing program of longitudinal studies guided by a comprehensive model of enrollment dynamics is vastly more useful than a series of non-cumulative one-shot efforts. While the conduct of individual cohort analyses or questionnaire studies may be little affected by the presence of such a model, the agenda of which studies to undertake at what time, as well as the manner of presenting complex results to decisionmakers are considerably helped by a sound and visible conceptual approach.

Naturally, there is no single recipe for success in constructing such an ongoing research program. Experience does, however, suggest a number of broad guidelines:

- Look at the "big picture" first. The most important single step in longitudinal enrollment research is to determine patterns of student flow for the entire institution. Estimating a cohort survival model can be of considerable help in forecasting enrollments and in guiding enrollment management policy. At the same time, focusing administrative attention on the "big picture" will help to communicate the results of more detailed studies by emphasizing how they relate to one another, and their implications for total enrollment.

- A partially estimated model is better than none. Establishing a longitudinal tracking database as described in this chapter may take time, and there is a strong temptation to wait to communicate results until a wide range of data are available. But even partial data may be of considerable value to enrollment man-

agement decisionmakers in ruling out less promising courses of action.

- Disaggregate enrollment until distinct behavioral patterns emerge. Aggregate models of student flow are useful, but can mask often considerable differences in enrollment dynamics typical of different groups of students. The best way of identifying such groups is through experimentation - by examining different successive cuts of enrollment statistics until behavioral groups that hang together emerge.

- Revisit the model continuously and check the validity of its assumptions. Longitudinal enroll ent models - even if grounded in considerable historical data - are only as good as the assumptions upon which they rest. These assumptions often change as a result of changes in policy or changes in the institution's operating environment. As a result, longitudinal enrollment models should never be seen as immutable. Rather they should be constantly reviewed in order to ensure their appropriateness, and to discover new insights into available student data.

REFERENCES

Ewell, P. T. (1984a). The self-regarding institution: Information for excellence. Boulder, CO: National Center for Higher Education Management Systems [NCHEMS].

Ewell, P. T. (1984b). Conducting student retention studies Boulder, CO: NCHEMS.

Hossler, D. (1984). Enrollment management: An integrated approach. New York: College Entrance Examination Board.

Kemerer, F. R., Baldridge, J. V., & Green, K. C. (1982). Strategies for effective enrollment management. Washington, DC: American Association of State Colleges and Universities.

Kraetsch, G. (1980, Winter). Methodology and limitations of Ohio enrollment projections. AIR Professional File (No. 11).

Kreutner, L., & Godfrey, E. S. (1981, Winter). Enrollment management - Vehicle for institutional and environmental control. The Board Review.

Lonabocker, L. & Halford, J. (1984, Fall). Projecting for the 80s: Managing student enrollments. College and University, 60 (1), 68-76.

NIE Study Group on the Conditions of Excellence in American Higher Education. (1984). Involvement in learning: Realizing the potential of American higher education. Washington, DC: U.S. Government Printing Office.

Stokey, E., & Zeckhauser, R. (1978). A primer for policy analysis. New York: W. W. Norton.

Wing, P. (1974). Higher education enrollment forecasting: A manual for state-level agencies. Boulder, CO: NCHEMS.

Zemsky, R., & Oedel, P. (1983). The structure of college choice. New York: College Entrance Examination Board

Zemsky, R. & Associates. (1980, Winter). Can colleges control enrollment? Educational Record, 10-15.

STUDYING STUDENT ATTRITION AND RETENTION

Patrick T. Terenzini
University of Georgia

Despite an enormous body of literature on the subject, and several excellent literature reviews (Spady, 1970; Cope & Hannah, 1975; Tinto, 1975; Noel, Levitz & Saluri, 1985), attrition and retention studies continue to be a common activity of institutional researchers as they and the administrators they serve seek to understand the dynamics of attrition on their own campuses and to develop ways of enhancing student retention. The reasons are not hard to discern: the supply of traditional high school graduates is shrinking and college-going rates (which grew steadily from 1950 to a high in 1976) are stabilizing, if not declining slightly (Kaufman, 1986). Colleges and universities across the country are facing or already experiencing enrollment drops, some of them substantial. At least some of the institutions that are maintaining enrollment levels may be doing so at the cost of reduced standards of admission or degree completion. Enrollment levels are, of course, intimately related to tuition at independent institutions and to state appropriations in public colleges and universities, and administrators are realizing that a penny saved is a penny earned: retaining a student who might otherwise withdraw means not having to recruit a replacement. The cost avoidance incentive applies elsewhere, as well: the time, energies, and resources required to register, advise, financially support, counsel, teach, and recreate a student who stays are the same as those for the student who eventually withdraws.

Moreover, retention rates are increasingly being used (appropriately or not) as one index of program or institutional quality and health. As will be seen below, retention information for program evaluation or for institutional self-study or review must be used cautiously. Such cautionary notes, however, will not entirely deflect the growing public pressure on postsecondary institutions to account for the effective and efficient use of public and private funds.

This chapter is intended to provide an introduction to the study of students' attendance behaviors on one's own campus. First, definitions and types of withdrawal are reviewed. Then, three basic designs for gathering data on why some students continue their enrollment while others withdraw are summarized and the assets and liabilities of each are briefly described. Finally, the chapter outlines the most common statistical procedures used to analyze the data assembled.

BASIC TERMS AND CONCEPTS

Before undertaking any study of attrition and retention, choices in three fundamental areas must be made: definitions, types, and purposes. These choices will have a direct bearing on the nature of any study's design, methods, and utility. Consider, for example, the apparently simple question: "What is the attrition/retention rate at Frostbite Falls State?" The response may seem more like an evasion than an answer.

- To how long a period of time does the question apply? Freshman to sophomore year? Four years? Five years? Ten years?
- If longer than a year, what about the student who returns after a term or two? Is that student a "dropout," even though that individual is currently enrolled?
- Are academic dismissals to be included, as well as those who withdrew voluntarily?
- Does the question refer to retention within each academic major program? (Probably not, but the question is neither unreasonable nor uninteresting.)
- What about the student who completes a course or two and leaves, but never intended to complete a degree, or even a full academic year?

Definitions

Definitional problems have bedevilled attrition and retention studies for decades. The problem lies in the fact that what constitutes "dropping out" is really a matter of perspective. It is important to be clear about this at the outset, for attrition is not necessarily a bad thing, something contrary to the best interests of institutions or students.

From the individual's point of view, withdrawal may indicate that the individual got from the institution exactly what he or she had come for. For these students, the definitions of "persistence" and "success" do not include completion of a conventional certificate or degree program, or even a full academic year. Similarly, not all withdrawals are attributable to a lack of ability. Where attrition is concerned, this belief may be one of the most prevalent - and pernicious - views of faculty members. Indeed, substantial evidence exists to indicate that students who withdraw commonly have higher scores on measures of academic aptitude than do continuing students. Students must accept (and we must give them) some responsibility for their own educations. Studies need to differentiate between students who withdraw after making a good-faith effort to succeed and those who withdraw in the face of self-induced failure. (See Pace, 1984, for guidance and

instruments for the evaluation of student effort.) Or, withdrawal may follow the student's discovery that the decision to attend a particular institution had been made for the wrong reasons or on poor information, and that educational or career goals will be better served at a different institution. Or, perhaps those educational or career goals have changed since matriculation and require attendance elsewhere. But if those changes were induced, at least in part, by the student's educational experiences at the present institution, then the withdrawal reflects personal and/or academic growth and constitutes both an individual and institutional achievement, not a failure.

From an institutional perspective, one must take into account institutional mission. For example, one of the primary functions of two-year institutions is to prepare students to transfer to four-year institutions. Such withdrawals should be sources of institutional and individual satisfaction in a mission accomplished, not causes for alarm. Indeed, in some vocational-technical schools, part of whose mission is to help students find employment, the institution may even encourage area employers to raid the classroom when looking for help. Finally, the decision to withdraw may be beyond the institution's power to influence. For example, a student may decide to withdraw significantly in advance of the actual withdrawal (little is known about the timing of such decisions). Similarly, the decision may be due to conditions beyond the institution's control (for example, problems at home).

While the matter of definition continues to be a subject of some debate, certain common categories appear to be emerging:

- PERSISTER: A person who is continuously enrolled (for the period of time under study). In certificate or degree-completion studies, this individual graduates "on time."

- STOPOUT: A student who leaves for a period of time, but who subsequently returns to continue study. This student graduates eventually.

- DROPOUT: A student who leaves and does not return during the period under study. (This student, of course, may subsequently prove to be a stopout.)

- ATTAINER: A student who leaves prior to certificate or degree completion, but after achieving a personal goal (for example, completion of a particular course, or acquisition of a particular skill).

This latter category is an important addition to our lexicon, recognizing the importance of students' initial intentions in the classification process. The "dropout" category in much of the earlier research, particularly that done at two-year institutions, is no doubt confounded by

the number of "attainers" who were included because of imprecise definitions.

Types of Withdrawal

Attrition may be of three general types. Internal attrition refers to the transfer of students to different major programs of study within the same institution. Such "attrition" may well be the most common form, but it is also the least threatening to an institution's financial condition and might be viewed as a normal consequence of the educational process. Internal attrition may, however, be a cause of concern to program or departmental administrators if it is very high, for their enrollments may influence the resources they receive from higher levels. Internal attrition in any given program or department becomes a cause of concern at the dean's or vice presidential level only if it is substantial in comparison with that of other majors.

Institutional attrition involves the student who leaves a particular college or university, even though that student may transfer to another institution with no interruption in schooling. For the original school, the student is gone, and the reasons for the leaving are worthy of study. This is the most frequently studied type of attrition, for it is directly related to institutional funding and is increasingly being used as an index of overall institutional health and quality.

Systemic attrition refers to students' withdrawal from all types of formal higher education. This type of attrition is a matter of interest at system, state, and federal levels because of its obvious public policy implications.

How Much Attrition Is It Reasonable to Expect?

Studies of student attrition and retention deal with one or more of four fundamental questions: 1) How many students are withdrawing? 2) When are students withdrawing? 3) Who is withdrawing? and 4) Why are they withdrawing? Peter Ewell, in his chapter in this volume, provides an excellent discussion of the data bases and procedures needed to answer the first three questions. This chapter will not repeat what is there, but, rather, focus in detail on the fourth question, partially addressed by Ewell. Before doing that, however, it may be helpful to know what the retention rates are at similar types of institutions: how much attrition is it reasonable to expect at one's institution? At what point on the continuum from perfect retention to complete exodus should an institution begin to be concerned about its own withdrawal rate?

Table 1 arrays national retention and graduation rates at different types of institutions over varying periods of time: the freshman year, two years, and graduation in three and five years. Three things are noteworthy about these figures: first (looking at the upper portion of the

table), the one-year retention rates for the three different cohorts of students are approximately the same, suggesting reasonable stability over time. While such stability may characterize the retention rates of each institutional type, considerable variability exists within each category.

TABLE 1

RETENTION AND GRADUATION BY TYPE OF INSTITUTION, NUMBER OF INSTITUTIONS REPORTING AND PERCENTAGE OF STUDENTS WHO ENTERED AS FRESHMEN

	Retention After One Year					
	1975-76		1976-77		1977-78	
Institution	N	%	N	%	N	%
2-Year Public	74	55	82	55	92	53
2-Year Private	27	63	29	64	30	63
Nonsectarian	12	63	12	62	12	64
Religious	15	61	17	65	18	62
4-Year Public	99	68	103	67	104	66
4-Year Private	207	71	222	71	227	71
Nonsectarian	66	73	72	73	76	74
Religious	141	71	150	70	151	68

	Retention After Two Years				Graduation			
	1975-77		1976-78		3 Years		5 Years	
Institution	N	%	N	%	N	%	N	%
2-Year Public					188	41		
2-Year Private					46	61		
Nonsectarian					18	63		
Religious					28	60		
4-Year Public	85	56	78	55			135	53
4-Year Private	176	57	178	57			306	60
Nonsectarian	52	63	55	63			105	64
Religious	124	55	123	54			201	58

NOTE: Reprinted from Beal and Noel (1979)

FROM: Lenning, O.T.; Beal, P.E., & Sauer, K. *Retention and Attrition: Evidence for Action and Research* Boulder: NCHEMS, 1980.

Second, notice the five-year graduation rates for four-year institutions (right-hand column of the lower table). While these figures may be somewhat higher than those from other sources, the point is clear: graduation four years after matriculation can hardly be considered the norm. While we used to believe that the student who failed to complete a degree in four years was atypical, in fact, an entering

freshman was about as likely to drop out as to complete a degree program over a four-year period. This continues to be the case.

Indeed, the national four-year degree retention rates have been virtually invariant for the last century. World War II caused a drop in the trend line, followed by a sharp but short-lived rise occasioned by the G.I. Bill, but with those exceptions, the national completion rate at baccalaureate degree granting institutions has held virtually constant at about 50 percent since 1880 (Tinto, 1982). It will be interesting to see over the next decade whether actions taken in the heat of the current national concern with retention will produce any appreciable change in that historical pattern.

The third point worth noting in Table 1 is that, while the tabled figures are for "retention," one can derive some interesting "attrition" statistics by subtracting the one-year retention figures (top table) from 100 percent. Having done that, one will see that, for both two- and four-year institutions, attrition is largely a first-year phenomenon. Using the public four-year institution figures to illustrate, one can see that on the average about 30 percent of our freshmen do not return for their sophomore year, compared with a five-year attrition rate of about 47 percent. Put another way, perhaps as much as 70 percent of those who will leave over a five-year period will withdraw before the second year (other sources suggest that figure may be closer to 50 percent, but the point remains the same). Among two-year institutions, the situation is even more dramatic: 46 percent attrition in the first year, compared to 59 percent attrition over a three-year period. That is, nearly 80 percent of all dropouts leave before the start of the second year. The message is clear: if one wishes to increase institutional retention rates, efforts should be concentrated on the first year.

Finding Out "Why?"

The most difficult and technically challenging question to answer is "Why do some students withdraw while others continue?" As Ewell suggests, much can be learned about the origins of attrition and retention on one's campus through analysis of longitudinal enrollment records. Groups that are particularly attrition-prone can be identified, and the "peak periods" of attrition can be delineated. Inferences based on such information are unavoidably limited, however, being restricted to the data elements contained in the longitudinal enrollment files. And if the research on the sources of attrition tells us anything, it is that the process is exceedingly complex, oftentimes more dependent upon what happens to students after they arrive on campus than on what they are like at the time of matriculation.

The number and variety of variables from which to choose in studying attrition on one's campus is, however, stag-

gering. Fortunately, several excellent sources are available to help one plan and organize for such a study. First, several excellent reviews of the attrition/retention literature are available. Those of Spady (1970) and Tinto (1975) are comprehensive and scholarly, but now also somewhat dated. Cope and Hannah (1975) offer a more readable summary, but this review is also now dated. Lenning (1982) provides an excellent discussion of variable selection and measurement issues and Noel, Levitz and Saluri (1985) provide a more current, but somewhat limited review.

Given the enormous range of variables from which to choose, some of which are better predictors of attrition than others, one will profit from consideration of several theoretical models of the attrition process. Such models can serve as valuable road maps in the selection of variables and in their causal ordering for study. Tinto's (1975) model is the best known and most frequently used in guiding research on student retention. Tinto views attrition and retention as a longitudinal process involving a complex series of socio-psychological interactions between the student and the institutional environment. According to this theory, the student brings to college such characteristics as family background, personal attributes, and pre-college school experiences, each of which is presumed to influence not only college performance, but also initial levels of goal and institutional commitment. These characteristics and commitments, in turn, interact with various structural and normative features of the particular college and lead to varying levels of integration into the academic and social systems of the institution. These levels of social and academic integration, in turn, influence subsequent levels or goal and institutional commitment. According to Tinto, "Other things being equal, the higher the degree of integration of the individual into the college systems, the greater will be (the) commitment to the specific institution and to the goal of college completion" (p. 96).

Bean (1980) has synthesized the research on student attrition and employee turnover in work organizations to devise a model that views the influences on students as analogs to those that lead to employee turnover. Anderson (1985) applied "force field analysis" to identify the influences on student achievement and retention. Whichever model one might adopt, the object is to arrive at some parsimonious selection of the major variables involved in student attrition and retention. Each of the models above attempts to guide such selection and can be used to save valuable time, energy, and space.

DESIGNING ATTRITION AND RETENTION STUDIES

Attrition/retention study development considerations fall into two general categories: research designs and analytical procedures. The former is concerned with how to

collect the necessary data, the latter with how to analyze the data.

Research Designs

Attrition study research designs may be of three general types: autopsy, cross-sectional, or longitudinal. In social science research, as in so many other things, "There is no such thing as a free lunch." Every design decision has its price, and it is important to know not only what the meal will be, but also what it will cost. This section, and the one following it, attempts to summarize what is gained and given away with the selection of each research design and analytical procedure. A more thorough discussion of these issues is given in Terenzini (1982), from which this is excerpted.

Autopsy designs, as the label implies, are ex post facto studies of attrition. The researcher identifies those students who have withdrawn during a specified period of time (for example, between freshman and sophomore year) and mails the former students a survey questionnaire asking about why they withdrew, how often they used certain academic and student services, their evaluations of those services, or other features of the college experience (for example, academic advising, living arrangement, major) and so on. Information on current activities is also usually solicited. The strength of this design lies in its focus on those who do not return to graduate. The design has considerable intuitive appeal: if one wishes to know why students are withdrawing, ask them! The design is also attractive because of its relatively low cost.

Logical and appealing as this approach may be, however, it is methodologically the weakest of those discussed here. The concept of controlled comparison is the foundation of scientific research, and as autopsy studies are usually carried out, students who continue their enrollment are not sampled and surveyed. Thus, while one might find from an autopsy survey that dropouts are highly critical of the academic advising they received, without a sample of continuing students (a control group), there is no valid or reliable basis for concluding that doing something about academic advising will reduce attrition: students who continue their enrollment may be equally critical of their advising. Even if continuing students are surveyed, information is likely to be collected at different times and under different circumstances than that for dropouts, thereby confounding comparisons.

Moreover, autopsy designs generally produce relatively low response rates, which are a threat to the generalizability of any findings. There is also a tendency for respondents to be more positive than nonrespondents in their attitudes toward the institution. It may well be that the responding "dropouts" are, in fact, "stopouts" who will eventually re-enroll. In addition, autopsy study data frequently do not include information that would be useful in controlling

background variables. It should be pointed out that nothing inherent in an autopsy design forces these methodological weaknesses. Rather, the weaknesses arise more from the researcher's lack of experience. Autopsy designs are probably most useful for obtaining anecdotal information.

Cross-sectional designs involve the one-time collection of data from currently enrolled students, typically near the end of a semester or academic year. A cross-sectional design produces an informational "snapshot" of the influences at a single point in time that are inclining students toward withdrawal or continued enrollment. Moreover, the design provides data from both persisters and dropouts (and stopouts and attainers), except, of course, that these groups cannot be differentiated one from another until the next term begins and the fact of enrollment or non-enrollment can be verified. At that point, the analytical groups are formed (stopouts will, of course, be confounded with dropouts at this stage), and the analysis of why students persisted or withdrew can begin, using the variables thought to influence attendance decisions and included in the survey questionnaire.

Certain design risks exist, however. Differences found between or among groups may be due to differences between or among the groups at the time they enrolled. Failure to consider such pre-college differences may lead the researcher to conclude (unwittingly and, perhaps, expensively) that the sources of attrition lie within the institution's control when, in fact, they do not. Students' admissions files will contain information for controlling some important pre-college differences (for example, high school achievement and academic aptitude), but information on other potentially confounding variables (for example, commitment to degree completion, educational and career aspirations, expectations of college) will not be available.

The strengths of this category of designs (relative to autopsy designs) include the likelihood of higher response rates, more reliable generalizability of findings to the entire student body, and, thus, greater utility of the information for other informational and evaluative purposes. In addition, this approach permits comparison of dropouts and persisters on the same measures, taken at the same time, and under similar conditions and at the very time they are presumably exerting their influence. Finally, cross-sectional designs are typically no more costly than autopsy designs. If a campus mail system is available for BOTH distribution and collection of questionnaires, it may even be less expensive than the autopsy design. (If the campus mail service does not handle the outgoing mail of students in residence halls, U.S. postage should be provided for the return of questionnaires. Do NOT assume students will drop off survey forms at some central location. They won't.)

The liabilities include the need for greater technical sophistication on the part of the researcher and more complex data-handling tasks (particularly if pre-matriculation differences are to be controlled), a longer time period required to complete the study (perhaps 6-9 months or more, less if studying semester-to-semester attrition), and somewhat higher costs.

Longitudinal designs are the most demanding, but also the most likely to produce valid information. If a cross-sectional design provides an informational snapshot of the influences on students' attendance behavior at one point in their college careers, longitudinal designs constitute something of a family album, marking changes in student attitudes and experiences over a period of time that may vary from a single semester or year to several years. Information is gathered at multiple points during that period. Typically, students are first surveyed before matriculation (for example, during a summer orientation session) on a variety of topics, including educational and family background, academic and career goals, aspirations and plans, expectations of college, and so on. Students are surveyed again at later points in time, perhaps at the end of each subsequent semester or year. Each follow-up data collection seeks to gather information on students' experiences during the year (or semester) just ending (for example, frequency of contact with faculty members, academic performance, peer relations, current attitudes, goals). The reliance of this approach (and of the two previous designs) on students' abilities to recall the year's (or semester's) experiences may be a source of error, however. For as noted earlier, the decision to withdraw may have been made well before it was actually acted upon. Because little is known about the typical time lapse between decision and action, this potential source of error is likely to remain until it is studied in greater detail.

At the start of the following year (or semester), registration records are checked to determine which students are "persisters" and which are not. Differences are sought between and among the groups in attitudes or experiences that might help explain why some students withdrew while others continued.

The Cooperative Institutional Research Program (of the Higher Education Research Institute at the University of California - Los Angeles), the Educational Testing Service (Princeton, NJ), the American College Testing Program (Iowa City, IA), and the National Center for Higher Education Management Systems (Boulder, CO) all provide instruments for surveying entering students and for following-up at later points in time.

Longitudinal designs provide extensive, planned control of confounding background variables, as well as more precise estimates of the institutional influences on attendance behavior. Such designs are the most internally valid

available for studying attrition and afford a measure of confidence in findings and associated conclusions that is not available with other designs. The price of such design rigor and confidence, however, is not trivial. Because of the decreasing number of respondents over time, large samples or a total census of entering students is needed. Handling large numbers of students, of course, increases the complexity of the data handling, the volume of materials to be handled, and the costs. Such designs also require at least one person on the study team who is familiar with longitudinal research methods and multivariate statistics. Such studies also require more time to complete, perhaps as much as fifteen months. Moreover, greater care is needed throughout the study: what was a minor oversight in an autopsy design can burgeon into major and costly problems in a longitudinal study. Considerable care and thought should be given to the selection and development of any design, for as will be seen below (if it is not already evident), decisions made about research designs have a direct and significant influence on everything that follows, including the analysis of the data.

ANALYTICAL PROCEDURES

At the same time that a study design is being selected (and absolutely before the data are collected), the institutional researcher must face a problem akin to that of a dog chasing a car: once the car is caught, what is to be done with it? Before the mid-1970s, the majority of attrition studies relied on univariate or bivariate statistical procedures. Univariate statistics include a variable's mean, median, mode and standard deviation and are derived to summarize a trait in a sample and to estimate the amount of that trait in a population. By themselves, they provide descriptive information about a group, but little else. Bivariate statistics (for example, chi-square tests of association or goodness-of-fit, t-tests, zero-order correlations, one-way analyses of variance) yield information about the relation between two variables in a sample, but they are ill-suited to the study of complex, multivariate phenomena, such as college attrition and retention.

Because one knows intuitively that students' decisions about continued college attendance are complex, survey questionnaires are designed to gather data on a large number of variables. Analysis of such data sets using bivariate procedures requires repeated calculation of the statistic each time one of the variables changes. Such repeated testing is subject to two problems. First, as the number of tests increases, so does the likelihood of rejecting an hypothesis of "no difference" when, in fact, no real difference exists (that is, the difference is due to chance). Out of every 100 such tests, one can expect five such errors. (That, after all, is the meaning of "$p < .05$.") The problem, of course, is that one has no way of knowing which of the differences are "true" and which

are "false" differences indicated by the significance tests.

The second problem with repeated bivariate tests is that the variables used in a retention study are, in all likelihood, correlated with one another. For example, high school achievement is related to academic aptitude, and both, in turn, are related to socio-economic status, and so on. Thus, if two or more variables are found through bivariate testing to be related to attrition, which of the two (or more) is the better (best) predictor of attrition? Once the first is taken into account, is the second still a useful predictor? Multiple bivariate analyses are not only unparsimonious, but also virtually uninterpretable with any degree of clarity. As Spady (1970, p. 77) has written, "further . . . bivariate research on the 'correlates' of dropping out should be abandoned. Now!"

One senses intuitively that attrition and retention behaviors are inherently multivariate - multiple variables in some as yet only poorly understood web of relations. Three types of multivariate statistical procedures are commonly used to help clarify that web: multiple regression, discriminant function analysis, and path analysis.

Multiple regression is a statistical procedure by means of which the contributions of multiple independent (predictor) variables to the prediction of a dependent variable (in this case, a dichotomous dependent variable: attrition or retention status) can be estimated both individually (the unique importance of each predictor variable) and collectively. Kim and Kohout (1975) offer a brief, relatively nontechnical introduction to the topic, and Kerlinger and Pedhazur (1973) present an extended and somewhat more mathematical discussion that is still readable, even for those who have a limited knowledge of statistics. Virtually all standard statistical computer packages have routines to perform multiple regressions.

Regression analyses may be of several forms, but the most frequently used in attrition analyses are overall or hierarchical. In an overall multiple regression, the contributions to prediction by all independent variables are tested simultaneously. The value of the R^2 statistic indicates the proportion of variance in the dependent variable (dropout v. persister status) that is attributable to all predictor variables acting simultaneously. That is, the R^2 indicates the power of the independent variables for explaining differences between dropouts and persisters. The magnitude of each variable's beta weight reflects the unique contribution of that variable to the predictive power of the regression equation, controlling for all other predictor variables. For example, if the beta weight of one variable is twice the magnitude of another, then the first variable contributes twice as much as the second to the explanation of attrition status. Such models do not shed a bright light on the importance of

institutionally-controllable variables on attrition, however, inasmuch as all variables - background and college experience variables - are tested simultaneously.

Hierarchical regression models can help isolate the institution's role, however. This type of regression permits the researcher to determine the unique predictive power of each variable, or set of variables, in a specified order. For example, the predictive power of all background variables can be tested first. Then, with the amount of variance in the dependent measure attributable to these variables known, the researcher can test the power of the college experience measures for explaining that proportion of the total attendance behavior variance not already explained by the background variables by adding to the original model the college experience variable and repeating the regression analysis. The incremental difference between the two R^2 statistics points to the "institutional contribution" to attrition, to that set of variables over which the institution has some control and which contribute to the understanding of attrition in ways not already explained by students' background traits.

Regression analyses have their limitations, however. First is the problem of multicollinearity, the correlations among multiple independent variables, which can seriously complicate interpretation of the regression coefficients. For example, when students' background traits are highly related to institutional experience variables, a hierarchical regression model will extract first the variance attributable to the background traits. The explanatory power of the institutional experience variables will, consequently, be reduced, producing reduced estimates of their influence and making determination of their true effect virtually impossible. When multicollinearity is high among the college experience variables, it is even more important to exercise caution in the interpretation of regression coefficients. Second, there is some debate over whether conventional, least-squares regression techniques are the best for analyzing dichotomous dependent variables. Some statisticians recommend probit models. A discussion of these models is, however, beyond the scope of this chapter. A third, and perhaps the most serious, drawback of both least-squares regression and probit analysis is their inability to handle more than two groups. As noted earlier in this chapter, withdrawal study categories may include as many as four groups (persisters, stopouts, dropouts and attainers), perhaps more, depending upon how they are formed. With more than two groups, other analytical techniques are needed.

Discriminant function analysis offers a solution. Discriminant analysis is something of a multiple-group extension of the two-group (that is, dichotomous dependent variable) multiple regression model. Indeed, the results of a regression analysis with a dichotomous dependent variable are directly proportional to those of a discriminant function analysis with two groups. The sta-

tistics produced are different in the two instances, but their interpretation would be the same. Discriminant analysis is used in retention studies to identify which of a set of multiple variables best differentiate between and among three or more groups. The point, here, is that in an attrition/retention study with more than two groups, discriminant analysis is the appropriate statistical procedure. All standard statistical computer packages (for example, SPSS, SAS) have routines for doing discriminant function analyses. Space does not permit more detailed description of discriminant analysis here, but readers are referred to Terenzini (1982) for a conceptual discussion of this analytical technique, or to Huberty (1975) or Tatsuoka (1971) for more thorough, but also more technical, discussions.

Path analysis, or structural equation modeling, has gained in popularity and frequency of use in attrition studies over the last five years. At one level, path analysis is simply the application of ordinary least-squares multiple regression techniques in a conceptually-structured way. Kerlinger and Pedhazur (1973) define path analysis as "a method for studying the direct and indirect effects of variables taken as causes [on] variables taken as effects. It is important to note that path analysis is not a method for discovering causes, but a method applied to a causal model formulated by the researcher on the basis of knowledge and theoretical considerations" (p. 305). The use of path analysis forces the researcher not only to identify the key variables, but also to hypothesize the causal relations among them that are believed to lead to, and culminate in, students' decisions to continue enrollment or to withdraw. Thus, the importance of having a theory: it guides and informs the construction of the path model.

CONCLUDING THOUGHTS

It is important to keep in mind throughout that attrition studies are, without exception, correlational. As such, no causal links between independent variables and students' decisions to withdraw or continue enrollment can be specified, even when using path analysis or other structural modelling technique (those procedures merely test causal relations hypothesized to exist). The finding that one or another variable is related to the withdrawal/continuation decision is, at best, a statement that the identified relation is probably not due to chance. It is not a statement of cause and effect.

As noted earlier, the design of attrition studies, like all social science studies, is a series of compromises: advantages gained in one quarter come at a price paid in another quarter. Methodological purity must be balanced against practical utility, cost, and timeliness. The trick is to optimize one's position, minimizing the investments of time and money without, in so doing, sacrificing so much methodological rigor as to make the study of questionable validity.

Finally, Ewell (in his chapter in this volume) stressed the importance of "an on-going program of longitudinal studies guided by a comprehensive model of enrollment dynamics." Such a program will depend upon no single approach. Each of the four questions outlined earlier (who, when, how many, and why) must be addressed if attrition is to be understood and effective retention programs designed.

REFERENCES

Anderson, E. (1985). Forces influencing student persistence and achievement. In L. Noel, R. Levitz, & D. Saluri (Eds.), Increasing student retention (pp. 44-61). San Francisco: Jossey-Bass.

Bean, J. P. (1980). Dropouts and turnover: The synthesis and test of a causal model of student attrition. Research in Higher Education, 12, 155-187.

Cope, R., & Hannah, W. (1975). Revolving college doors: The causes and consequences of dropping out, stopping out, and transferring. New York: Wiley-Interscience.

Huberty, C. J. (1975). Discriminant analysis. Review of Educational Research, 45, 543-598.

Kaufman, N.S. (1986). Shifts in postsecondary clientele. In P. M. Callan (Issue Ed.), Environmental scanning for strategic leadership. New Directions for Institutional Research (No. 52). San Francisco: Jossey-Bass.

Kerlinger, F. N., & Pedhazur, E. J. (1973). Multiple regression in behavioral research. New York: Holt Rinehart and Winston.

Kim, J., & Kohout, F. J. (1975). Multiple regression analysis: Subprogram regression. In N. H. Nie, C. H. Hull, J. G. Jenkins, K. Steinbrenner, & D. H. Bent (Eds.) SPSS: Statistical package for the social sciences (2nd ed.). New York: McGraw-Hill.

Lenning, O.T. (1982). Variable-selection and measurement concerns. In E. T. Pascarella (Issue Ed.), Studying student attrition. New Directions for Institutional Research: (No. 36; pp. 35-53). San Francisco: Jossey-Bass.

Noel, L., Levitz, R., & Saluri, D. (1985). Increasing student retention. San Francisco: Jossey-Bass.

Pace, C. R. (1984). Measuring the quality of college student experiences. Los Angeles: University of California at Los Angeles Graduate School of Education, Higher Education Research Institute.

Spady, W. (1970). Dropouts from higher education: An interdisciplinary review and synthesis. Interchange, 1, 109-121.

Tatsuoka, M. M. (1971). Multivariate analysis: Techniques for educational and psychological research. New York: Wiley.

Terenzini, P. T. (1982). Designing attrition studies. In E. T. Pascarella (Issue Ed.), Studying student attrition. New Directions for Institutional Research. (No. 36, pp. 55-71). San Francisco: Jossey-Bass.

Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. Review of Educational Research, 45, 89-125.

Tinto, V. (1982). Limits of theory and practice in student retention. Journal of Higher Education, 53 687-700.

Tinto, V. (1987). Leaving college. Chicago: University of Chicago Press.

DESIGNING AND CONDUCTING NEEDS ASSESSMENT STUDIES

Mary Kathryne Wilders
Moraine Valley Community College

Most colleges and postsecondary educational institutions are beginning to recognize that the educational environment in which they served the public in the 1970s and early 1980s no longer exists and that many new problems and perspectives need to be addressed. This situation has caused administrators at many institutions to increase their demand for reliable, projectable educational data. As a result, many college administrators and planners are now interested in conducting needs assessment studies to provide their institutions with this educational data.

NEEDS, WANTS, AND DEMANDS

Before examining educational needs assessment, the difference between needs, wants, and demands must be made clear. The clarification of these terms will make the design and use of a needs assessment instrument easier.

What people actually need may be quite different from their wants and/or demands. Needs are a state of felt deprivation of some basic satisfaction. They cannot be created.

According to Abraham Maslow, human needs can be placed in a hierarchy of order: physiological, safety, social, esteem, and self-actualization (Maslow, 1954). People act to satisfy their basic needs before satisfying their higher needs.

Human wants are desires for specific satisfiers of these deeper needs. Wants will vary from individual to individual based upon life experience and individual taste. For example, a person who wants to be a systems analyst may want to go to a postsecondary institution for training. A high school senior interested in pursuing further education may want to attend a small private college with a strong academic program and a reputation for turning out a high percentage of graduates who are subsequently selected for graduate fellowships.

Unlike needs, wants can be created. Human wants are continually shaped and reshaped by social forces, technological innovation, and the desire for change (Kotler & Fox, 1985).

Demands are wants for specific products or services that are backed up by an ability and willingness to purchase them. A college that wants to be on the forefront of training computer technologists must have the money to purchase the latest computer equipment.

Educational institutions often confuse wants and needs. A college may think a student needs a course in design

drafting but the student is really in need of a job. The student who finds another college that provides a better training program will have a new want but the same need.

NEEDS ASSESSMENT IN EDUCATION

Needs assessment has many definitions in the field of education. Most often needs assessment in education takes the form of a discrepancy approach that attempts to measure the gap between some desired or acceptable condition and the actual or perceived condition. The most comprehensive definition in literature related to needs assessment appears to be that of Robert Kaufman. He uses a discrepancy approach and defines educational needs assessment as a means of measuring the "gap between What Is and What Should Be in terms of results" (Kaufman, 1982).

This process involves a formal analysis that shows the gaps between current results and desired results. The gaps are then arranged in priority order and the needs to be resolved are then selected. The result of a needs assessment may show too little or too much of a condition or state of affairs. Therefore, Kaufman also views needs assessment as a tool for problem identification and justification. This tool requires consensus in planning and setting the priorities of needs (Kaufman & English, 1979).

A MODEL FOR CONDUCTING A NEEDS ASSESSMENT

Needs assessment involves many steps, such as identifying and clarifying who, why, what, how, and when issues. In order to maximize the planning for a needs assessment, several questions must be answered.

Who wants the needs assessment? It is important to delineate at the onset of the process exactly whose needs are of concern. Two possible areas of focus are the needs within the institution and the needs external to the institution. Needs assessments that focus within the institution may include, for example, questions about a change in enrollment patterns, course and program revisions, pressure for organizational change, or desire for system change. Needs assessments that have a focus external to the institution may result from forces such as prospective students, change in requirements for funding, accrediting agencies, community pressure, or state pressure.

Why is a needs assessment requested? This is an important area to consider when conducting a needs assessment. The assessor should be wary of possible unstated reasons for conducting a needs assessment. Is there a hidden agenda? Perhaps the reason for conducting the study is to legitimize what the college is already doing. Are political reasons motivating the study? Perhaps this study is just part of a ritual of soliciting opinions before taking an unpopular action, such as program cutbacks. Is the purpose of the study to raise the consciousness of the target au-

dience involved? Will key clientele be involved in the study? Is the needs assessment study supported by the president of the college? Based on the answers to these questions, the purposes, goals and objectives of the needs assessment can be written to focus on providing the required information.

What is the scope of the needs assessment? The scope of a study deals with both its content and the number and type of respondents sought. Each of these dimensions may be narrow or broad in scope. A needs assessment with a narrow scope will be concentrated, specific, and detailed. For instance, a study covering only developmental programs in reading or mathematics (content) or a study of freshman success (respondents) would be narrow in scope. A needs assessment study with a broad scope will be general and diffused. For instance, a study of community interest in continuing education (respondents) or the study of all instructional components (content) is broad in scope.

The resources available, both human and financial, will also determine the scope of the study. The cost of producing questionnaires, hiring interviewers, and the use of the computer, will have an impact on the scope of the study. The timeframe in which the study is to be conducted and the personnel available will also influence the scope of the study.

Whose needs are to be studied and at what level? A comprehensive, two-level classification of groups and entities of possible concern in assessment of needs was developed by Oscar Lenning and others as part of an NCHEMS Outcome Structure (Lenning et al, 1978). The first level of classification contains individual/group clients; interest-based communities, geographic-based communities; aggregate of people; and other audiences.

How should needs assessment be conducted? It is not only important to determine the type and amount of data that should be collected for the study but also the method to be used in the collection of data. The constraints on data collection, and analyzing, interpreting, and using data should also be examined.

TYPES OF DATA

An issue to be addressed in implementing a needs assessment is the type and amount of data to be collected. Should qualitative or quantitative data be used or should it be a combination of both types of data? Keeping in mind the difference between wants and needs will help the assessor distinguish between what must be collected and what would be nice to have. Trend analysis of demographic data, resource inventories, causal analysis, social indicators, group processes, performance measures, self-reports, historic data, futures data, and census data are some examples of measures that can be collected. Each has an associated

cost and benefit. The needs and resources of the institution will jointly determine which are to be utilized.

Collecting valid and reliable evidence of needs is a crucial part of every needs assessment study. A number of relevant indicators and measures usually apply, and some factors may be better than others. For instance, a factor such as whose needs are being assessed can affect the validity of the indicator or measure. Therefore, multiple indicators and measures should be used whenever possible. Using multiple measures assures that the need is real. Both subjective and objective data should also be used whenever this is possible.

DATA DEVELOPMENT METHODS

In the remainder of this section five ways to develop data will be examined: surveys, social indicators, group processes, futures methods, and causal analyses.

Surveys should be administered to a sample of the members of the group(s) under study. A survey may also be disseminated to other relevant groups that indirectly have knowledge about the group under study. These questionnaires can be written or recorded interviews. The instruments will help the assessor gather opinions, preferences, and perceptions of fact. (This is the procedure most widely used in the assessment of needs. Many studies use both the questionnaire and the interview. The interview is used as a follow-up device to ensure a higher rate of response or as a means of probing certain questions in depth.)

The most effective type of needs assessment survey asks respondents for informed opinion based on either personal expertise and knowledge, or for facts about themselves or others about which they have direct knowledge (Witkin, 1984). Never should direct questions about needs be asked, since this will result in a list of preferences, desires, or wants. The survey should also gather information about what is and what should be.

If the survey is not the sole source of data, then it should be determined exactly when in the process it should be used. If the survey is administered at the beginning of the study, then it can be used for exploring ideas that will be analyzed later in more detail by other methods of data collection. If the survey is administered at the end of the study, it can be used to corroborate or refute the needs identified previously by other means. The timing of the survey will dictate the content of the questions and the depth to which they should be explored.

Needs surveys use questions that ask for opinions regarding the degree to which a condition should hold. The perceived performance of objectives, the degree of agreement with a set of conditions, the frequency in which a behavior occurs, or the degree of satisfaction with processes within

a system can also be determined by a survey. The level of difficulty that the service receivers have with specified tasks, the preference for a program or activity, demographic information, past and present use of services, and other factual data can also be collected via a survey (Witkin, 1984).

Response formats for questionnaires typically use scales that have an odd number of points (usually five) and are anchored with descriptors at each point or at the ends. The response format may be either single-response or multiple-response. Single-response questions are generally used to elicit the respondent's perception of the degree or extent of need in specified areas. Multiple-response questions usually center on what is and what should be, thus making discrepancy analysis easier.

Other techniques used to design questions are forced-choice answers that establish the strength or intensity of judgments; the budget-allocation method that distributes a fixed number of points over a set of items resulting in a priority ranking; and the critical incident technique which supplies indicators or exemplars of the attainment of goals. Open-ended questions should be used sparingly, since there are many variables involved in interpreting the data.

Surveys are often less costly and more time efficient than interactive methods. There are many advantages of using a survey. Large amounts of data can be gathered in a short time span; a wide geographic area can be covered; there is less chance for sidetracking and irrelevant inputs; the process is easy to manage; and the results can usually be computerized. The disadvantages include the possibility of respondents misunderstanding questions; a potential for low return rates; and failure to take into account cultural and/or linguistic differences in respondent groups.

Social indicators are demographic, descriptive, and statistical in nature. Some examples of social indicators that have been used in educational needs assessment studies include goal indicators; test and performance data; indicators related to people, programs, and organizational variables; and indicators of training needs. Data of this type are most useful in identifying the current status of a curriculum, faculty, or program.

Data that identify the size and characteristics of the population groups with particular needs, the symptoms of those needs, and the scope of the problem are essential. These indicators cannot determine the needs of groups, but when used with other information, they can show evidence of need. Data of this type already exist in various forms and in various locations, although it may be gathered by means of an interview. Census data is readily available and provides information on the demographics of the population.

Rossi and Gilmartin (1980) propose a conceptual framework for the establishment of a time-series data base that can be used to identify social indicators. They suggest the following steps for constructing a time-series data base: develop a conceptual framework; identify and interrelate variables to be monitored; identify and screen existing data sources; identify and screen statistics from the data sources; prepare a social-indicator report; and perform a time-series analysis.

Houston and others (1978) propose that social indicators be collected on people, programs, and the organization. Variables suggested for the people indicators are student variables (personal characteristics and attitudes, behavior, knowledge); professor variables (personal characteristics, competence, attitudes toward students); administrator variables (personal characteristics, attitude, competence); and parents and community variables (attitude toward college, priorities, goals).

Variables suggested for programs include curricula (content, sequence, resources); lesson, unit, and module (content and sequence, strategies and methods). For organizations, variables include governance (policies, composition); administration (personnel, students, instruction); and management climate (college satisfaction, interpersonal relations).

Procedures using a group process in needs assessment center around public hearings and/or forums, group discussions, focus groups and the modified Delphi technique.

Groups can be gathered to validate existing data or to provide new information. Although questionnaires provide the same information using written responses, group processes have the added advantage of active public involvement of key groups of individuals. Another advantage of a group process is that the respondent can ask for clarification of questions, resulting in an adjustment in methodology to match the communication styles of different ethnic and linguistic groups.

The disadvantages associated with group processes center around the selection of participants; the necessity for skilled leadership; the structure and quality of the thinking of the group; time and costs involved training interviewers; time necessary to plan, schedule, and carry out the process; possible distortion of data inputs and interpretation by leaders or interviewers; reluctance of people to participate in a group process; the tendency for "me too" responses in the group; and greater difficulty in the interpretation and the use of the results. Careful advance planning on the part of the needs assessment project leader can alleviate some of the above mentioned problems.

When a needs assessment is conducted for long-range planning (more than three years into the future), futures

methods should be utilized. Some techniques of futures forecasting are scenarios, Delphi, and cross-impact analysis.

A scenario is a story about a possible future. The basis for this story is a set of plausible, integrated events or conditions in society. The Delphi technique is a series of questionnaires or small group processes that help achieve consensus of opinion among experts. Cross-impact analysis uses a matrix to determine the impact of one group of factors on another group of factors. These techniques are fraught with problems such as who should be consulted about possible futures? How are data from the various participants summarized? Is the future based on past trends? It is difficult to predict the future, and these techniques should never be used as a sole source of information for needs assessment.

When trying to identify factors underlying why the need has not been met, causal analysis methods should be used. This method bridges the gap from the identification of needs to the development of solutions. The most common technique used is fault tree analysis. This methodology uses logic diagramming to relate combinations of possible events, or subsystems within a system, to show how they interact to produce a predefined undesired event. A thorough explanation of this procedure is too complex for this paper; however, a good reference is Fault Tree Analysis: A Research Tool for Educational Planning by B. R. Witkin and K. G. Stephens (1968).

Reliability and validity are the primary criteria to be considered when making a selection of which data are to be used. Consideration should be given to data requirements that allow little flexibility and data that are readily available. In addition, the appropriateness of the analytical procedures and tests planned, ease of scoring and tabulation, collection and analysis costs and whether administrators and planners will be able to see the relationship of the data relative to their practical concerns should be considered. Standards of quality, amounts of data, data analysis, and balancing costs against requirements are additional constraints to consider.

ANALYZING, INTERPRETING, AND USING DATA

Interpretation and the use of needs data are also crucial elements in a needs assessment study. A needs assessment study should not only identify needs, but rank them according to how critical they are. Methods for setting priorities should be built into the needs assessment from the onset of the study. Clear criteria for analyzing and converting the data to priorities, using decision rules and systematic procedures to establish levels of need should be determined early in the planning phase of a needs assessment study.

In addition, other criteria should be considered, such as the probability of successful resolution of the problem or need; the availability of time, personnel, and facilities for solving the need; the monies necessary to meet the need; and the number of individuals directly affected by meeting the need. The managerial and political arenas also contribute important criteria when deciding which needs to address. Considering these criteria during the needs assessment study facilitates the realistic identification of alternatives to solve the need. Furthermore, some of the managerial problems mentioned by the respondents may be alleviated before the needs assessment is completed. If the needs assessment addresses the political and managerial concerns, the utility of the results will be enhanced.

STRENGTHS AND WEAKNESSES OF NEEDS ASSESSMENT

A variety of approaches to needs assessment have been tried; each approach has areas of strength and areas of weakness. One of the greatest advantages of a needs assessment is the involvement of consumers in the needs assessment process. A representative sample of the general public is given the opportunity to impact directly on the future.

The value of the information obtained outweighs all other advantages. Needs assessment provides decisionmakers with an extensive data base from which programs can be developed in an effort to meet the needs of specific audiences within their jurisdiction more effectively. Additionally, once this data base is established it can then be used, along with longitudinal research, to evaluate the program's effectiveness in meeting the identified constituent needs.

Witkin (1984), on the other hand, questions whether needs assessment helps the decisionmaker in making decisions concerning program plans and allocation of resources, or whether it just helps the decisionmaker organize the information.

Other questions can be raised about the use of discrepancy scores as the only basis for ranking needs. Needs assessments often fail to provide insightful and specific information about needs since decisionmakers receive information that is not easily translated into a course of action. Often what is done under the guise of a needs assessment is actually a measure of attitude.

Some needs assessment approaches only try to identify needs without ranking the needs or trying to determine why the needs occurred. Often researchers attempt to rank needs through the use of simple decision rules that do not consider enough factors or that consider each factor in isolation from the others. Such rules are seldom effective. Another problem with many of the needs assessment models is that they focus on current goals and objectives and provide no information for long-range planning.

SUMMARY

A needs assessment effort should not generate grand general statements about groups under observation or the individuals who belong to these groups. Results of a needs assessment can only be valued in terms of the accumulation of detailed facts that illuminate and support a decision. However, needs assessment, when properly planned and used, is clearly a viable tool to assist administrators and faculty members concerned about meeting client and community needs.

REFERENCES

Babbi, E. R. (1973). Survey research methods. Belmont, CA: Wadsworth Publishing Company, Inc.

Houston, W. R., & others. (1978). Assessing school/college/community needs. Omaha: The Center for Urban Education, The University of Nebraska at Omaha.

Kaufman, R. (1982). Identifying and solving problems: A system approach. (3rd ed.) San Diego, CA: University Associates.

Kaufman, R., & English, F. W. (1979). Needs assessment: Concept and application. Englewood Cliffs, NJ: Educational Technology Publications.

Kotler, P., & Fox, F. A. (1985). Strategic marketing for educational institutions. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Lenning, O., & others. (1978). A structure for the outcomes of post-secondary education. Boulder, CO: National Center for Higher Education Management Systems.

Maslow, A. H. (1954). Motivation and personality. New York: Harper & Row.

Rossi, R. J., & Gilmartin, K. J. (1980). The handbook of social indicators: Sources, characteristics, and analysis. New York: Garland STPM Press.

Sudman, S., & Bradburn, N. M. (1982). Asking questions: A practical guide to Questionnaire Design. San Francisco: Jossey-Bass.

Witkin, B. R. (1984). Assessing needs in educational and social programs. San Francisco: Jossey-Bass.

Witkin, B. R., & Stevens, K. G. (1968). Fault tree analysis: A research tool for educational planning. ERIC Document No. ED 081 039.

PROGRAM EVALUATION IN HIGHER EDUCATION

Richard F. Wilson
University of Illinois at Urbana-Champaign

INTRODUCTION

Interest in program evaluation in higher education is sustained by a number of forces, but one of the most abiding is financial distress. When resources are limited, institutional leaders seek information that will help them understand which programs are of highest quality and should be protected, which ones have good potential and should be encouraged, and which ones are weak and should be curtailed or terminated. On the surface this seems like a relatively straightforward task, and many program evaluation processes have been launched out of a concern about limited resources.

When resources are more plentiful, the sustaining force often is a desire to make things better - to invest in a program so that it is strengthened. Those in leadership positions want to know where resources can be spent to have maximum impact in terms of improving programs.

A third major force behind program evaluation is what has been termed external accountability. The advent of state-level coordinating and governing boards and of consumerism in higher education, as well as the interest in value-added education, have all played a role in increasing the pressure on institutions to demonstrate to external constituencies that resources are being used wisely and effectively.

At least one of these three forces is present on most campuses all of the time. Consequently, the question is not whether program evaluation should occur but how it should be done.

The terms program evaluation and program review are frequently used interchangeably and refer to those activities in which judgments are made about initial program performance and priorities for the future (commonly known as formative evaluation) or about the longer term success of a program in reaching established goals or performance standards (commonly referred to as a summative evaluation). An overly simplistic way to think about the difference between these two types of program evaluations is that a formative evaluation is focused on how to make a program better while a summative evaluation is more concerned with whether a program has been successful or effective.

CURRENT LEVEL OF ACTIVITY

Program evaluations can be initiated at any organizational level, but most frequently occur at the institutional, system, or state levels. In some states, program evalu-

ation occurs at all three levels. Barak studied the prevalence of program review at these different levels and found extensive activity. Based on a survey of a stratified sample of institutions, he found that over 82 percent had some type of program review process. In the same study he reported that such reviews were undertaken in half of the multicampus systems. Finally, he reported that all fifty states have some type of process for reviewing existing programs in public institutions and a few include private institutions as well (Barak, 1982).

The point is that the programs of some institutions are subject to review on at least three different levels. Although there are some unique aspects to these reviews, there is also a fair amount of overlap and duplication. Wallhaus (1982) has tried to clarify the program review domains of institutions versus state-level agencies and Craven (1980) has tried to carve out a unique program review role for multicampus systems, but considerable overlap remains. Few institutions have the luxury of designing a program review process that does not have to be coordinated in some way with review initiatives at other organizational levels.

The key steps in a program review process are presented in Table 1. Although each of these steps is important, special attention will be given to identifying purposes (Step 1), defining the scope of the review process and selecting the evaluative approach (Step 2), collecting data (Step 4), and linking the results of the evaluation to other decisionmaking processes (Step 7). The implementation of these steps has important ramifications for those providing institutional research services.

PURPOSES

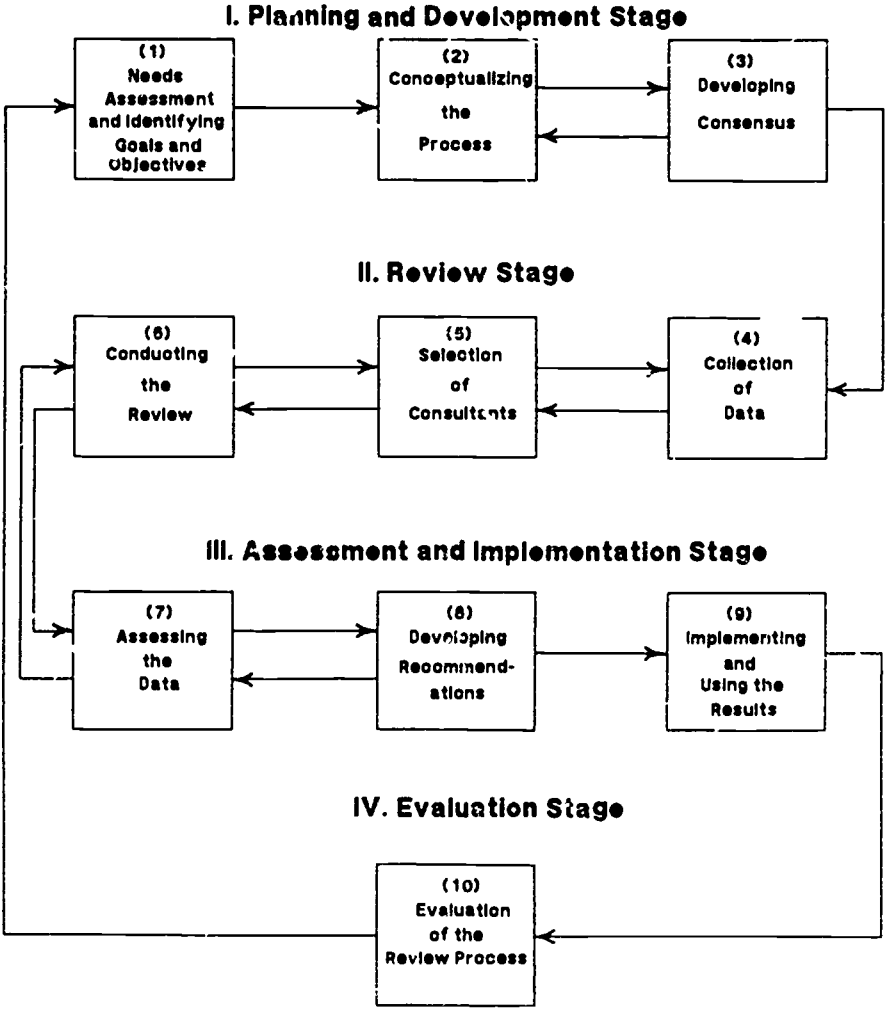
There are at least five major reasons for establishing a program review process:

1. to help programs improve,
2. to meet multicampus or state-level review mandates,
3. to demonstrate institutional responsiveness to external constituencies,
4. to provide a basis for allocating and reallocating resources,
5. to provide information to be used in making decisions about program discontinuance.

It is not unusual for an institutional review process to embrace several of these purposes. In fact, it may be imperative to do so if there is to be any hope of building support for the process. Different groups will want to accomplish different objectives. In spite of such accommodations, most institutions have one purpose that is more important than the others. Establishing priority among purposes is crucial to the success of a review process. It is very difficult to collect data and make judgments

Table 1

Suggested Stages and Steps in Developing a Program Review System



Source: Barak, 1982, p. 70.

that are equally responsive to a number of different purposes.

SCOPE OF REVIEWS

An early issue that must be addressed is the selection of programs to be included in the review process. It is safe to assume that all academic programs will be included; what is not clear is whether to include academic support units (for example, the library), administrative units (for example, the institutional research office), and research and public service units. Although these units are quite disparate (thus requiring a unique evaluation plan each time one is done), there is mounting pressure to evaluate units of this type. At most institutions, the faculty have tolerated initial attention being focused on academic programs, but have been asking that other campus units be reviewed as well. A key issue is whether academic and non-academic units should be reviewed with the same process.

APPROACHES

Once there is agreement on the purposes of the program review, the next series of questions relate to what criteria should be used, what data should be collected, who should be consulted, who should make the evaluative judgments, and who should receive the results. The answers to these questions depend upon the approach to evaluation. Four popular approaches will be discussed below.

The goal-based approach is the most popular evaluation strategy in higher education. This approach draws upon the work of Tyler (1949) and emphasizes the identification of program goals and objectives, the specification of performance standards that relate to these goals and objectives, and the collection of data that will enable judgments to be made about whether performance expectations have been met. The widespread use of this approach stems largely from its orderly and rational nature.

Stake (1975) and Guba and Lincoln (1981) have played key roles in the development of the responsive approach to evaluation. This approach minimizes the attention to formal goals and objectives, focusing instead on the issues and concerns of those who have some stake in the program under review. There are no preconceived ideas about what should be studied in a responsive evaluation; the issues are defined through a series of interviews or surveys with a program's constituencies. Data are collected relative to these issues and perceptions. In essence, an attempt is made to determine what a program is actually accomplishing and what issues need attention, regardless of the formally stated goals and objectives.

In contrast to the two approaches discussed above, the decisionmaking approach to evaluation emphasizes the linkage between evaluation activities and the information

needs of those in decisionmaking roles. There is less interest in the formal goals and objectives of a program and in the concerns of constituencies than in securing information that will enable wise decisions to be made. There are no preconceived notions about how data are to be collected and analyzed; what is done is dependent upon the decision that needs to be made. Stufflebeam (1971) has been a leading proponent of the decisionmaking approach and developed the CIPP model which, in effect, organized evaluations around decisions relating to Context, Input. Process, and Product.

Eisner (1975) and House (1978) have given considerable attention to the connoisseurship approach to evaluation. In higher education this model takes the form of peer reviews, commonly referred to as outside review teams. The essence of this approach is the value placed on the experience and insight of the expert in the field. That individual's stature and knowledge is viewed as crucial to a deep understanding of what is being accomplished. Data may be collected from a variety of sources, but in the end it is the connoisseur who organizes and analyzes all of the data and provides a report that is based on perceptions of the strengths and weaknesses of a program.

The four approaches discussed above do not constitute an exhaustive list. There are variations on each one and some others that have been proposed. However, these four have discernible differences and have received the most attention and use.

COLLECTION OF DATA

Evaluation is as much an art as it is a science. There are few instances where a question can be answered or an issue resolved with a single piece of data. One searches for insights that will increase the confidence that a correct judgment is being made. These insights are more informed when data from a variety of sources point in the same general direction. If the task is to assess the quality of instruction in a program, perceptions may be secured from currently enrolled students, but this information is not sufficient. Faculty members may need to visit classes, the success of graduates may need to be monitored, the change in student performance over time may need to be measured, and alumni reactions may need to be secured. Data may also be collected from other institutions with comparable programs. If the information from these sources points in the same direction, those performing the evaluation can have greater confidence that a correct judgment is being made.

LINK TO DECISIONMAKING

A question that needs careful attention is how the results of evaluations are linked to decisionmaking processes. Evaluations are undertaken for a variety of reasons, but there is an implicit assumption that the results of such

efforts will improve decisionmaking. The problem is that evaluation processes tend to take on a life of their own (perhaps because they frequently take so much time) and are not connected in a meaningful way to such things as planning and budgeting. It may not be crucial for evaluations to drive these other processes, but it is important that there be a link. Otherwise, results will not be used and enthusiasm for evaluation will decline.

KEY PRINCIPLES FOR SUCCESSFUL REVIEWS

There are several principles that appear crucial to successful review processes. These principles apply to all stages of a review and will be discussed briefly below.

Fairness. Although it may seem patently obvious, the one overriding concern of everyone involved in evaluation is whether the process is fair and whether those involved are treated fairly. It is easy to endorse the concept of fairness, but this commitment sometimes gets overlooked in the rush to complete an evaluation. The way in which an evaluation is conducted is often attacked as much as the results.

Timeliness. It is very difficult to sustain interest in an evaluation or to convince people that the effort is worthwhile if it takes too much time. There is a tendency to design evaluations that are very elaborate. There is nothing wrong with an elaborate evaluation as long as the resources (people and money) are available to complete the task in a timely way. Otherwise, evaluation efforts will dissipate energies and be counterproductive.

Responsiveness. An institution may have a standard evaluation process, but there probably will never be a standard evaluation. Every program is somewhat different, and the issues worth exploring will certainly vary. The credibility of the evaluation will depend on the evaluators' willingness to adapt procedures to unique circumstances. One of the benefits of microcomputers is that surveys that once had to be standardized can now be altered with minimal effort.

SIGNIFICANT ISSUES

Three issues have been selected for special attention because of the frequency with which they emerge in program evaluation and because of their importance to the success of these processes. The issues are (1) the adoption of multiple purposes; (2) the assessment of quality; and (3) the use of results.

Multiple Purposes. Most institutions and agencies go through a rather lengthy process of consultation in establishing a program review system. This effort usually results in the identification of a number of purposes that may be served. There will be those who see such reviews as an opportunity to identify ways in which programs can

be strengthened, and others who will feel that the chief contribution will be to ascertain where cuts can be made so that funds can be reallocated to more pressing needs. In some instances, there may be little enthusiasm for establishing a review process other than to satisfy requirements of system-level or state-level agencies.

In order to build the necessary consensus to initiate program reviews, it is not unusual for all of these expectations to get folded together. Although this sounds reasonable, there are serious questions about whether such an amalgamation can work. For example, those responsible for implementing a review process will find it difficult to know how to deal with a weak program if they are charged to provide advice on how programs can be improved and at the same time are asked to provide advice on program discontinuance and resource reallocation. In such a situation, those in programs under review may be confused about how forthright they can be with the reviewers. Can they be completely open, even sharing perceptions based on limited information in hopes that the program reviewers will be able to help with a potential problem, or must interactions be restrained, dealing only with documented facts, because some aspect of the program, or perhaps its very existence, is at stake?

Similarly, problems can arise over the conflicting expectations of institutions, system offices, and state-level agencies. It is not that any of these expectations are unreasonable, it is simply difficult sometimes to combine them in the same review process. For example, an institution may decide that it is desirable to retain a program dealing with an esoteric subject because of that program's significance to the development of knowledge even though enrollments may be dwindling or non-existent. At another level this program may appear unjustifiable, constituting a misuse of resources. Often this kind of debate ensues whenever student enrollment is used to the exclusion of other information as an indicator of value or worth.

It is important, therefore, that purposes be chosen carefully and that the relationship of the institutional review process to those conducted at other levels be examined to ensure that review expectations are compatible. Where such compatibility does not exist, it may be best for all concerned for separate review processes to be established. It will be difficult to sustain a review process where program expectations conflict. To eliminate such conflicts, some institutions have assigned program improvement to one review committee and budget review to another.

Assessing Quality. For those engaged in program review, there is no more elusive task than the assessment of quality; yet such assessments are fundamental to the review process. The irony is that people believe that they can "see" when quality is present (even better when it is not present) even though no one has found a good way to measure it. Frequently, what has been seen is hard to verify.

Because of this situation, a lot of attention has been devoted to defining the concept and to identifying measures that will provide some insight into whether quality is present.

There are four views of quality that have been discussed widely. These views and some examples of indicators for each one are displayed in Table 2.

Table 2

VIEWS OF QUALITY AND REPRESENTATIVE INDICATORS

Reputational View

- Peer judgements of the quality of program, students, faculty, or resources

Resources View

- Student selectivity
- Student demand
- Faculty prestige
- Faculty training
- Faculty teaching loads
- Budget affluence
- Library holdings
- Equipment adequacy
- Size of endowment

Outcomes View

- Faculty scholarly productivity
- Faculty awards and honors
- Faculty research support
- Faculty teaching performance
- Student achievement following graduation
- Student placement
- Student achievement
- Alumni satisfaction

Value-Added View

- Change in students' cognitive abilities
- Student personal development
- Student career development
- Social benefits

Source: Conrad and Wilson, 1985, p 51

The reputational view traces its roots to the connoisseurship model of evaluation and places heavy reliance on the informed judgment of experts. In higher education this takes the form of national surveys of faculty members or department heads who are presumed to be knowledgeable about the work of their colleagues at other institutions. A recent example of a study of this type was undertaken by Jones, Lindzey, and Coggeshall (1982).

There is an intuitive appeal to reputational studies because most academics believe that quality can be perceived even if it cannot be measured. Also, there is a belief that the human mind is capable of assimilating a broad range of information and transforming it into a judgment about quality in a way that cannot be matched. On the other hand, there are critics who have studied the rankings and found them lacking in several respects. The time dependency of the rankings, the inadequate knowledge of the raters about many of the programs being ranked, and the correlation between ratings and size of faculty are examples of such criticisms (Conrad & Blackburn, 1985).

One of the more traditional ways of assessing quality is termed the resources view. Essentially, one looks at the human, financial, or material resources available to a program or institution to determine its quality. This view of quality relies on such measures as the test scores of entering students, the institutions where faculty members obtained their degrees, and the number of books in the library. These data are relatively easy to obtain and certainly have some bearing on the quality of a program. Conversely, these measures tell you more about the quality of the resources with which the program must work than about what the program has contributed. Also, it is very difficult to identify the unique resources of a program; many resources, for example, the faculty, are shared by several programs.

The outcomes view of quality moves to the other end of the educational process and focuses on measures of results. Typical measures include number of faculty publications, student performance on standardized tests, and studies of alumni satisfaction. The significance of the outcomes view in terms of assessing quality is that attention is focused on accomplishments of those who have been a part of a program. The major limitation of this view is that it is difficult to isolate the contribution that the program made to whatever performance is measured (Astin, 1980). For example, if an institution enrolls a bright student and four years later that student scores well on an achievement test, how much of the test performance can be attributed to the program that the student has completed as opposed to the student's own ability?

The final conception of quality is termed the value-added view because of the attempt to address the program contribution problem of the outcomes view. Those espousing the value-added approach to assessing quality focus on

measuring change over time. A typical strategy is to have students take an achievement test upon entry and at graduation, thus determining what value the institution has added. A major problem with this view of quality is that it is frequently difficult for those in a program to agree precisely on what a student should accomplish while enrolled. This lack of agreement leads to the development of very general objectives, for example, the ability to exercise independent judgment, that are almost impossible to measure. There also is a danger that attempts to determine the value added will focus on the measurable, thus trivializing the educational process. The value-added view of quality is very attractive but has limitations that should be recognized.

This discussion demonstrates that the assessment of quality is difficult. There are no easy solutions. For most institutions this means that all four views of quality are integrated into the program review process. By collecting data from a variety of sources, those responsible for program review hope to gain some insight into the quality of what is being done.

Use of Evaluations. It should not be surprising that the use of evaluations is of paramount importance to all those involved in program review. Considerable time and attention is devoted to evaluation tasks and everyone is interested in whether the effort is worthwhile. Like most aspects of evaluation, however, this issue is not easy to resolve.

If evaluations were implemented immediately, then it should be possible to determine which recommendations had been implemented and which ones had not, thus keeping a scorecard on the process. Unfortunately, this is not the fate of most evaluations. On most campuses an evaluation report is one piece of information for those in decisionmaking roles. This information gets combined with other data before a decision is made. The point is that the assessment of use of evaluations must focus on the longer term as well as the more immediate consequences of an evaluation. Braskamp and Brown (1980) have also cautioned that it is a mistake to focus on direct and observable consequences of evaluations; evaluations frequently have subtle, catalytic, or unintended effects that must be considered.

Because of the complexity of the task, research on use of evaluations is restricted to a few case studies of evaluation efforts at the institutional and state agency levels (Mingle, 1978; Smith, 1979). Studies of utilization are needed in order to clarify what is being accomplished. Until better evidence is produced, those responsible for program review processes will labor under the suspicion that most evaluations are not very useful.

CHALLENGES FOR INSTITUTIONAL RESEARCH OFFICES

In many colleges and universities, the institutional research office is instrumental to program review efforts. At a very fundamental level, those making evaluative judgments need data that those in institutional research capacities can provide. This historical function will persist, but creative thought must be given to ways of displaying data so that they are meaningful even to the occasional users. When those responsible for an evaluation ask for data, the response should include attention not only to the data need but also to the presentation need.

A second and related challenge is to devote attention to collecting and analyzing data that will provide further insight into program quality. Much of the data currently in use is assumed to be important for such judgments, but research on the validity of these measures is limited. The task is to identify new data sources and validate their usefulness, to undertake research on data currently in use to determine their value, and to synthesize existing research on data that have been studied extensively. A special task in this regard relates to assessing the contribution of data that purport to measure "valued-added" components of a program.

Finally, institutional research needs to be done on the use of evaluation results. Those with evaluation responsibilities are probably not in the best position to conduct this research. Institutional researchers can help. Much work needs to be done in terms of specifying a method for determining use and collecting meaningful data.

REFERENCES

Astin, A. W. (1980). When does a college deserve to be called "high quality"? In Improving teaching and institutional quality. Current Issues in Higher Education (No. 1). Washington, DC: American Association for Higher Education.

Barak, R. J. (1982). Program review in higher education: Within and without. Boulder, CO: National Center for Higher Education Management Systems.

Braskamp, L. A., & Brown, R. D. (1980). Editors' Notes. In L. A. Braskamp & R. D. Brown (Issue Eds.), Utilization of evaluative information. New Directions for Program Evaluation (No. 5). San Francisco: Jossey-Bass.

Conrad, C. F., & Blackburn, R. T. (1985). Research on program quality: A review and critique of the literature. J. C. Smart (Ed.), Higher education: Handbook of theory and research, (Vol. 1). New York: Agathon Press.

Conrad, C. F., & Wilson, R. F. (1985). Academic program reviews. ASHE-ERIC Higher Education Research Report, (No. 5). Washington, DC: Association for the Study of Higher Education.

Craven, E. C. (1980). Evaluating program performance. In P. Jedamus & M. W. Peterson (Eds.), Improving Academic Management, San Francisco: Jossey-Bass.

Eisner, E. W. (1975). The perceptive eye: Toward the reformation of educational evaluation. Stanford, CA: Stanford Evaluation Consortium.

Guba, E. G., & Lincoln, Y. S. (1981). Effective evaluation: Improving the usefulness of evaluation results through responsive and naturalistic approaches. San Francisco: Jossey-Bass.

House, E. R. (1978). Asumptions underlying evaluation models. Educational Researcher, 7, 4-12.

Jones, L. V., Lindzey, G., & Coggeshall, P. E. (Eds.) (1982). An assessment of research-doctorate in the United States. (5 Vols.). Washington, DC: National Academy Press.

Mingle, J. R. (1978). Influencing academic outcomes: The power and impact of statewide program review. In C. Fincher, W. T. Furniss, J. R.Mingle, & D. S. Spence (Eds.), The closing system of academic employment. Atlanta: Southern Regional Education Board.

Smith, S. (1979). Program review: How much can we expect? Unpublished manuscript. University of California-Berkeley.

Stake, R. E. (Ed.) (1975). Evaluating the arts in education: A responsive approach. Columbus, OH: Merrill.

Stufflebeam, D. L., et al. (1971). Educational evaluation and decision making. Itasca, IL: Peacock.

Tyler, R. W. (1949). Basic principles of curriculum and instruction: Syllabus for education 360. Chicago: University of Chicago Press.

Wallhaus, R. A. (1982). Process issues in state-level program reviews. In R. E. Wilson (Issue Ed.), Designing academic program reviews, New Directions for Higher Education (No. 37). San Francisco: Jossey-Bass.

BUDGETING AND FINANCIAL PLANNING

Stefan D. Bloomfield
Oregon State University

In the last decade, the subject of budgeting and financial planning has evolved from a little understood and somewhat mysterious process often jealously guarded as a prerogative of the Vice President for Finance or the Budget Director, to a major institutional activity involving the President, other vice presidents, and often many administrative and academic offices. Part of this evolution reflects the financial strains besetting many institutions of higher education - both private and public. Part also reflects the growing awareness of the many ways in which financial decisions can determine academic policy and other institutional choices - and represents an attempt to remedy inappropriate "tail wagging the dog" situations. Finally, part mirrors the increasingly sophisticated managerial skills that many higher education administrators are now bringing to their positions, and their recognition that financial planning is a key determinant of all areas of institutional activity.

As the procedures of budgeting and financial planning have become more analytical and have included consideration of more of the institution's diverse activities, institutional researchers have become drawn into these activities - both as a central source for a wide range of institutional data, and as quantitatively oriented professionals able to help refine existing budgeting and financial planning procedures. The growing importance of these activities is indicated by the special attention given to this topic by the relevant professional organizations. Budgeting and financial planning is the topic of numerous seminars, conference sessions, and journal articles sponsored by the Association for Institutional Research, the National Association of College and University Business Officers (NACUBO), the Society for College and University Planning (SCUP), as well as a variety of organizations catering to upper-level college and university administrators. This chapter presents an overview of the concepts and terminology of budgeting and financial planning that should be part of the basic toolkit of every institutional research and planning professional.

THE VOCABULARY OF FUND ACCOUNTING

An introduction to budgeting and financial planning in higher education must, of necessity, start with a discussion of the special vocabulary used by practitioners in this field. The necessity for this special vocabulary arises in part because higher education, like many other public service enterprises, differs in a very fundamental way from business organizations: its resources are marshalled not to create a profit that will increase the wealth of its owners, but rather to carry out a mission

accruing to the betterment of society as a whole. This difference is reflected in the respective financial documents: the financial documents of the business sector are designed to display expenditures and resulting revenues in such a way as to show how effective such expenditures were in bringing about the desired accumulation of net profit. The financial reports of higher education, by contrast, are designed to display the receipt and expenditure of resources on behalf of the major service activities of the institution, thereby rendering to governing bodies and various funding sources an accounting of the institution's stewardship of such resources.

The method adopted for this purpose within higher education is that of fund accounting, whose specific procedures are continually being refined by NACUBO. Because the resources provided to a college or university may carry a variety of restrictions or other special stipulations, separate funds are established to account for the use of such resources. Each fund itself is an independent accounting entity with a separate tracking of assets, liabilities, and a resulting fund balance. Financial planning and budgeting practices in higher education usually mirror the structure of these fund accounts, making a knowledge of this system an essential prerequisite for the institutional research and planning professional.

As a start toward this goal, the following is a summary of the classifications and other terminology of higher education fund accounting, as abstracted from NACUBO's authoritative manual, College and University Business Administration (Welzenbach, 1982).

FUND GROUPS

One of the most important distinctions encountered in fund accounting is that between restricted and unrestricted funds. Restricted funds contain resources that are provided to the institution subject to legally binding restrictions on their use imposed by a donor or funding agency external to the institution. These funds - such as grants or contracts for specific projects, operating funds for specific programs, appropriated funds for specific purchases, or donations to particular departments - must be used for the specified purpose only. The unrestricted funds have no such limited use predetermined by the provider, although their purpose may subsequently be limited by action of the institution's governing authority.

Most funds of the institution belong to one of six major fund groups (of which the first is most important for our purposes):

1. Current funds: These are the operating funds of the institution, to be expended in the near term for its activities of teaching, research, and service. Current funds will generally include both restricted and unrestricted funds, with the unrestricted funds carrying

special importance as the institution's primary source of flexibility for budget shifts and resource reallocation.

2. Loan funds: These are funds made available for loans to students, faculty, and staff. They may have been given to the institution specifically for these purposes, or may be otherwise unrestricted current funds designated by the governing authority for this use.

3. Endowment and similar funds: These are funds provided to the institution under the stipulation that the principal is nonexpendable. The income from such funds can be restricted by the provider, or may be usable as unrestricted current funds. Quasi-endowment is the name for funds that the institution's governing authority has chosen to set aside for the moment as if they were an endowment.

4. Annuity and life income funds: These constitute assets received by the institution under an agreement by which the donors in exchange will receive an annuity or other income for the remainder of their lives.

5. Plant funds: These are funds relating to the institution's physical plant. They include not only the capital assets of the institution, but also account for debt servicing of such assets and their renewal and replacement. This category, however, does not include ongoing operation and maintenance of the physical plant (which is accounted for as part of the current funds).

6. Agency funds: These are funds that the institution holds as "banker" for other entities of the university community, whether individual students and faculty, or organizations such as student associations, living groups, sports clubs, etc.

Revenues

The day-to-day operations of the institution are supported by the unrestricted and restricted revenues coming into the various current funds. Because an understanding of the sources of these revenues is critical to effective budgeting and financial planning, a detailed categorization of revenues has been established to aid in their analysis. The major sources of current fund revenues are the following:

1. Tuition and fees: These revenues include all tuition and fees assessed to students for current educational operations. For accounting purposes a student's tuition and fees are included in this category even when there is no intention to collect them from the student - so tuition waivers and scholarships are shown as actual revenues to this category (but then also appear as off-setting expenses in other accounts).

2. Government appropriations: The most important of these revenues for public institutions are the state or local appropriations that commonly fund the majority of the institution's current operations. Such revenues may derive from legislative acts or from local taxing decisions, and are predominately unrestricted when intended for current operations - although some portion may be restricted in nature. (Governmental revenues are generally considered unrestricted if changes in their use can be made without legislative approval.) For comparative studies of public institutions, the analyst should know whether the institutions involved are required to remit their tuition and fees to the state as an offset against state appropriations.

3. Government grants and contracts: Government grants and contracts expended for current operations are most often restricted in use, but the indirect cost recovery from such grants is unrestricted revenue that commonly may be applied to any area of current operations.

4. Private gifts, grants, and contracts: These revenues from nongovernmental sources include unrestricted gifts and grants as well as restricted gifts, contracts, and grants expended for current operations. As with governmental grants and contracts, any indirect cost recovery from such grants and contracts is classified as unrestricted income.

5. Endowment income: Any income from the institution's endowment that is expended for current operations, whether restricted or unrestricted, becomes part of the institution's current funds.

6. Sales and services: Any income from the rental or sale of goods or services in conjunction with the institution's educational, research, or public service missions is included as part of current fund revenues.

Expenditures

Current fund expenditures include all costs incurred for goods and services used in the conduct of the institution's operations. In the same way that current fund revenues are categorized to identify their source and the conditions of their receipt, so are current fund expenditures grouped so that their use can be planned, monitored, and accounted for. For our purposes, the two most important groupings of current funds expenditures are the following:

1. Education and General: This is the category of expenses (often referred to as E&G expenses) most commonly examined when analyzing or comparing an institution's "cost of instruction". The subcategories of Education and General expense are fairly well standardized throughout higher education, and should be well understood by all institutional research and planning professionals. They consist of the following:

a. Instruction: These expenditures for instructional activities include faculty and staff salaries and all departmental operating expenses except for academic administration at the dean's level.

b. Research: This category includes all separately budgeted research activities. Research efforts funded on the department's own budget, however, are included as part of the expenses for instruction.

c. Public service: This category includes noninstructional services rendered to individuals and organizations external to the institution.

d. Academic support: These are the expenses incurred for direct support of the institution's prime missions. Included are such items as libraries and museums, instructional media services and academic computing support, as well as academic administration at the dean's level.

e. Student services: This category includes the expenses of the offices providing direct support to students (which may include the registrar, admissions, financial aid, counseling, etc.), as well as the provision of social and cultural activities for students.

f. Institutional support: This category includes the expenses associated with executive administration of the institution, as well as central administrative operations such as the business office, personnel services, administrative computing, planning and budgeting, and public relations.

g. Operation and maintenance of plant: This category includes expenses for the operation and maintenance of the institution's physical plant, including custodial services, utilities, and grounds maintenance. It does not include major remodeling or renovation.

h. Scholarships and Fellowships: This category, representing a substantial body of funds mostly "passed through" the institution, is often excluded from E&G comparisons aimed at determining the costs required to run an institution of higher education.

2. Auxiliary Enterprises: These are the various units of the institution that are managed on a self-support (cost recovery) basis - charging students, faculty, or staff for the services being provided. Such units are expected to "break-even" in both public and private institutions. Common examples include housing and food service operations, parking, intercollegiate athletics, and some student health services.

BUDGETING AND FINANCIAL PLANNING

Although planning is commonly viewed as an activity focusing on the future, its success depends on an insightful understanding of the historical evolution of the institution and recognition of the relevant internal and external forces currently at play. This is especially true in the area of budgeting and financial planning, as reflected in the four activities that need to be carried on continuously and simultaneously as part of a modern, integrated budgeting and planning process.

1. Systematic and ad hoc analyses of the most recently completed budget period, with the goal of identifying (with benefit of hindsight) particular strengths or weaknesses in the previous budgeting process.

2. Continuous oversight of the current period's budget, monitoring for significant deviations from forecasted revenues or expenditures, and seeking an understanding of their causes.

3. Development of the next period's budget, building in new institutional resource allocations or reallocations as permitted on the basis of short-term expense and revenue forecasts.

4. Updating the institution's moderate-term (3 to 5 year) financial plan, incorporating anticipated new initiatives as deemed feasible in light of apparent trends in the economic, demographic, political, and technological environment.

It is not unusual for these complementary activities to be assigned to different administrative offices: the Business Office may monitor the current budget, the Budget Office may be developing next year's budget, and longer-range financial forecasting may be done in a vice presidential suite. The institutional research function, however, can usefully enter into all aspects of budgeting and financial planning, and can be especially helpful in bringing an overall, integrative perspective to these activities.

Financial Forecasting

Attempts to place budgeting and financial planning on a more "scientific" basis have emphasized the identification of driving factors for each of the different types of revenues and expenses described in the preceding pages. The analyst is asked to identify particular factors - some economic, some political, some emotional, but most outside the institution's direct control - that "drive" or induce future changes. An understanding of these factors then forms the basis for a forecast of the likely direction and extent of such changes. Comprehensive financial planning requires such an analysis of each of the major sources of revenue and items of expense. For many elements of ex-

pense, this process reduces to a relatively straightforward estimation of various future rates of inflation in materials and salaries. Other expense forecasts, however, require assessments of possible technological changes (such as in the computerization of campuses) or political forces (such as compliance with new civil rights requirements).

A similar analysis of driving factors is then performed for the elements of revenue. This latter process, however, can be much more complex and speculative, involving such tasks as: estimating future levels of state appropriations based on assessment of the state's economic health and political climate, estimating future tuition revenues based on knowledge of demographics and the institution's recruiting plans, estimating future gift revenues based on knowledge of evolving tax laws and the institution's development plans, and estimating future grant and contract revenues based on an assessment of the changing funding climate at sponsoring agencies and the institution's specific research strengths. These considerations lead ultimately to construction of a scenario upon which the institution's total revenues and expenditures can be estimated. The incorporation of such assessments into a formal analytical model is the subject of another rapidly evolving branch of applied science: mathematical financial planning models (see, for example, Hopkins & Massy, 1981; Wyatt, Emery, & Landis, 1979).

Public vs. Private Institutions

The process of estimating future revenue flows and expense requirements highlights some of the ways in which budgeting and financial planning in private institutions differ from that in publicly supported institutions. Some differences are obvious. In the area of tuition revenues, for example, private institutions have wide-ranging freedom in setting tuition levels and admissions policies, but are precariously vulnerable to the vagaries of the marketplace. Public institutions often have no voice in such decisions, but generally can depend on known E&G revenues from appropriated or voted funds.

As regards gift revenues, until recently this source of funds was received in large amounts almost exclusively by private institutions, which in some cases have been able to incorporate substantial endowment income into their financial plans and operating budgets. Now many public institutions have mounted aggressive fund-raising programs as well, so this element of financial planning is ceasing to be a distinguishing factor between the two sectors.

Less obvious differences between the public and private sector in higher education budgeting and financial planning often appear on the expense side of the ledger. Bureaucratic regulations imposed on public institutions can severely restrict flexibility in virtually all areas of expense - from mandated salary and wage guidelines (par-

ticularly for support staff), to purchasing regulations, restrictions on carry-over of funds and transfer of funds between expense categories, and control of capital construction decisions. To the extent that such managerial decisions remain within the prerogative of private institutions, their budgetary and financial planning procedures can take full advantage of this added flexibility. Thus, although the necessity for, and the basic definitional elements of, budgeting and financial planning are common to all institutions of higher education, the specific analyses required (and the institutional research support requested) can vary dramatically between private and public institutions.

CAPITAL BUDGETING

Two areas of expenditure often given specialized treatment in budgeting and financial planning are capital budgeting with respect to both the institution's physical plant and its equipment inventory. The large expenditures associated with construction and major remodeling of buildings clearly require extensive financial analysis beyond that needed for most other budgetary items (Callnan & Collins, 1986). For public institutions of higher education the capital construction budget is often considered through a process of legislative deliberation and appropriations completely separate from that of the E&G budget. The capital construction budget for private institutions is also often subject to a separate consideration process, but in this case such analyses must more explicitly assess the possible effects of the proposed capital construction budget on the institution's operating budget (see, for example, Harran, 1985). Note that private institutions are often aided in capital projects by state guarantees of bonds or pooling of bonds, and by loans and other assistance from the federal government. In both the public and private sectors, these capital budgeting decisions are characterized by detailed financial assessments, which may involve questions of bonding authority, estimates of securities markets, development of nontraditional "creative" financing schemes, and other technical considerations clearly setting such decisions apart from the less specialized analysis given to the operating budget.

Another somewhat specialized budgeting and allocation problem is the distribution of equipment monies to the various departments and operating units of the institution. These resources are used to replace existing equipment worn out from extended service or made obsolete by new technological advances, as well as to add new equipment to fill gaps in current inventories or to facilitate entry into emerging instructional or research areas. Particularly in the area of scientific equipment, the decreasing rate of new building construction and cutbacks in federal research outlays have reduced some of the traditional sources of new equipment funds. As a result, the burden of equipment acquisition, maintenance, and replacement has shifted increasingly onto the institution's operating

budget. In both the laboratory and the office, this increasing reliance on operating budgets comes at a time of unparalleled acceleration in the rate of equipment obsolescence on purely technological grounds. A method of analyzing the budgetary impact of an institutional equipment replacement and acquisition program is presented in Bloomfield and Groll (in press).

CONCLUSION

Budgeting and financial planning continue to be rapidly developing areas of institutional interest and concern, responding to both internal developments and external events. Public institutions are especially vulnerable to externally imposed changes in budgeting and planning processes as states change their mandated procedures, formats, or philosophy in this area. Such changes may occur as states experiment with zero-based budgeting or a variety of formula budgeting techniques, or as different incentive financing strategies are proposed (Leslie, 1984; Spence & Weathersby, 1981). Institutional budgeting processes are equally being reshaped by changing concerns and emerging practices within the institution itself - one of the most significant being the advent of strategic planning. Some of these issues are explored in Berg and Skogley (1985).

The common thread through all these aspects of higher education budgeting and financial planning, however, is the need for a thorough knowledge of the institution and an appreciation and understanding of the external forces affecting the institution. This is a role tailor-made for the institutional research and planning professional, and one that will become increasingly important as these financial activities continue to assume a central role in institutional administration.

REFERENCES

Berg, D., & Skogley, G. (Issue Eds.) (1985). Making the budget process work. New Directions for Institutional Research, (No. 52). San Francisco: Jossey-Bass.

Bloomfield, S., & Groll, B., (1987). An equipment resource allocation model. Unpublished manuscript, Oregon State University.

Callnan, M., & Collins, J. (1986). A clear challenge for the future: Capital asset management and planning at colleges and universities. Planning for Higher Education, 14(3), 30-38.

Harran, S. (1985). Capital budgeting: A small college system. In R. Fecher (Issue Ed.), Applying corporate management strategies. New Directions in Higher Education, No. 50. San Francisco: Jossey-Bass.

Hopkins, D., & Massy, W. (1981). Planning models for colleges and universities. Stanford, CA: Stanford University Press.

Leslie, L. (Ed.) (1984). Responding to new realities in funding. New Directions in Institutional Research, No. 43. San Francisco: Jossey-Bass.

Spence, D., & Weathersby, G. (1981). Changing patterns of state funding. In J. Mingle & Associates, Challenges of retrenchment. San Francisco: Jossey-Bass.

Welzenbach, I. (Ed.) (1982). College and university business administration, Washington, DC: National Association of College and University Business Officers.

Wyatt, J., Emery J., & Landis, C. (Eds.) (1979). Financial planning models: Concepts and case studies in colleges and universities. Princeton, NJ: EDUCOM.

ECONOMIC IMPACT STUDIES

Mary K. Kinnick
Portland State University

R. Dan Walleri
Mt. Hood Community College

INTRODUCTION

Institutional researchers have been conducting economic impact studies in increasing numbers since the mid-60s (Salley, 1977; Erwin & Miller, 1982). These studies are part of the larger framework of studying the outcomes or impacts of higher education, although economic impact studies measure outcomes incidental to the primary mission or purposes of higher education.

A major objective of these studies is to find out "whether it costs a community more or less than it gains economically by having a college or university in its midst" (Caffrey & Isaacs, 1971, p. ix). Those conducting the studies typically adopt a fairly narrow operational definition of economic impact: the additional money, in the form of income and jobs, in circulation within a local area, region or state, which can be credited to the existence of the college(s) or university(ies).

In the widely used input-output models developed by Caffrey and Isaacs (1971), the college is regarded as an export sector of the local economy (Salley, 1977). The amount of funds originating from outside the district or other locale and spent by the college, its faculty and students, in the district or locale is identified, and estimates of total impact are derived.

Two important considerations in doing an economic impact study are understanding the conceptual and methodological issues and deciding, specifically, "how-to-do-it". The focus here is on the former for several reasons. There is no one best way to conduct these studies. Local circumstances vary greatly, and ample material is available describing how these studies have been structured and conducted by particular institutions and systems. Less attention in the published literature has been given to conceptual and methodological issues. The aim, then, is to identify and discuss some of the issues that will face those considering the conduct of an economic impact study, including deciding to conduct one in the first place.

A word of caution at the outset. These studies can be highly demanding of time and energy. And, by implication, choosing this area of study over another suggests it is considered more important. It is of particular concern that exclusive attention may be given to this one kind of impact or outcome in the interest of a perceived short-term public relations payoff. Such an approach may prove det-

rimental in the longer-term, resulting from neglect of the outcomes stemming from the primary mission of colleges and universities - higher education. If conducted, an economic impact study should be made part of a larger effort to study outcomes related to the broader educational and social/cultural purposes of higher education (Bowen, 1977).

OVERVIEW AND RECENT DEVELOPMENTS

Table 1 represents an impressionistic overview of the scope and focus of economic impact studies. The Caffrey-Isaacs models were developed for use by single institutions attempting to estimate their economic impact on the local area. There have been no major alternative models or approaches developed since their classic work. Their models, or derivations of them, continue to be used whether or not they are strictly appropriate. Adaptations of these models have been developed by those conducting state-level studies (Erwin & Miller, 1982). Focus has ranged from one on money in circulation, to multiple impacts, to the emerging and broader focus on economic development.

Table 1

Focus and Scope of Economic Impact Studies

Focus	Scope		
	Single Institution	System	Statewide
	---Number of Studies*---		
Money in circulation	many	some	few
Multiple economic impacts	some	some	few
Economic development	few	few	few

* estimated from a scanning of several available bibliographies

While no one best source of information on conducting studies on the impact to local or statewide economic development has been identified an excellent place to begin is a recent publication of the American Association of State Colleges and Universities (SRI International, 1986). In the future, increased attention will probably be given to conducting studies of impact on economic development and less to the more traditional and narrower studies of economic impact. This move should be encouraged since most conceptual frameworks of economic development include attention to the broader and more fundamental roles of higher education. The SRI report (p.13) for instance, points to

roles in increasing accessible technology, developing a skilled and flexible workforce, entrepreneurial management, and increasing the availability of risk capital.

DECIDING TO CONDUCT A STUDY

The major motive for conducting economic impact studies is political. With scarce resources and many more players competing for public and private dollars, especially by those in the not-for-profit sector, many believe it to be in their interest to demonstrate the economic return on local or state investments, particularly in terms of additional income and jobs. Economic impact studies are one way of determining the extent to which higher education adds dollar value to the local or state economy.

Primary audiences for the results vary by kind of institution (for example, community college, private college, state university, etc.). Most study sponsors, however, target state legislatures, the business community and taxpayers, and current and prospective funders. Little research is available about what difference these studies make, positively or negatively. There is a strong belief, however, that they can help show that higher education is not a drain on local or state resources, but, rather, a stimulus. Having the results may not produce measurable gains; but not having the information may limit the ability of the institution to compete effectively with others for funds and other kinds of support.

Some of the potential advantages and disadvantages associated with conducting these studies are shown in Table 2. Results should be, and typically are, part of a more comprehensive promotional and marketing effort of the institution or system. Before initiating a study, then, it is critical that the study clearly and deliberately be made part of the strategy of top leadership for developing support for the institution. Before committing to an economic impact study, three questions should be posed and discussed: Will such a study help develop support for the institution? Is the capability and expertise to conduct such a study available? and, To whom and how are the study results to be disseminated?

Table 2

Potential Advantages and Liabilities
of Conducting an Economic Impact Study

Potential Advantages
• Inform the public that higher education does not operate in a financial vacuum.
• Demonstrate that higher education makes a positive contribution to the economy.
• Influence attitudes of business leaders, using language they understand - dollars and cents - about the value of higher education.
• Influence voters and legislators to continue support.
Potential Liabilities
• Risk suggesting to the public that economic impact is a central mission of higher education.
• Identify college expenditure data which may stir up new questions and critics.
• Risk appearance of the study as a self-serving effort.
• Risk negative consequences if the study is found to be conceptually or procedurally flawed.
• Cost of conducting the study, including the opportunity cost of forgoing other studies.

STUDY APPROACHES

The Basic Model

The predominant model cited by those conducting these studies continues to be Caffrey-Isaacs, based on the "export sector" assumption described earlier in this chapter. A major limitation of the model is its scope of application. That is, the model assumes that a significant proportion of the institution's enrollment is drawn from outside the local area, and thus, that this enrollment represents an influx of new money into the local economy.

Despite this limitation, community colleges and other institutions serving largely local populations use the model, selecting those components which do appear appropriate and modifying others. Urban universities and community colleges should be encouraged to include in the cash flow

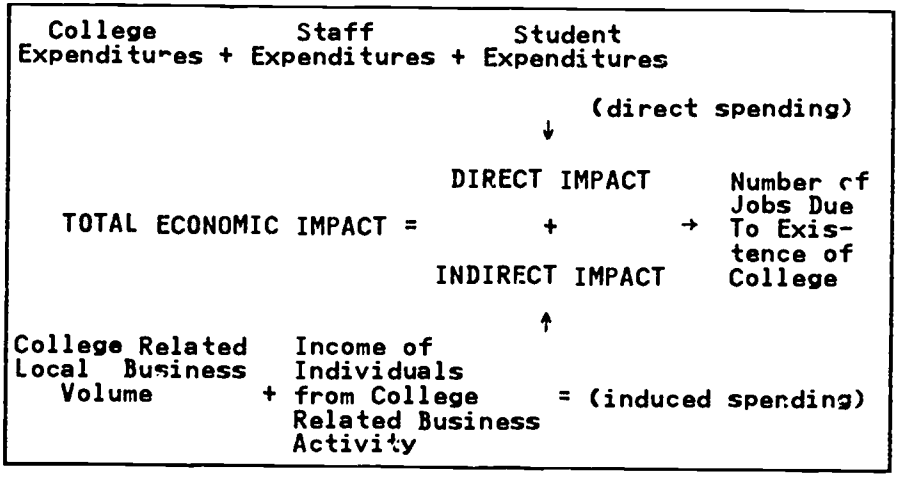
traced only dollars spent locally by students who indicate they would move to another geographic area to attend school were the college not in existence.

There are no major alternatives to the Caffrey-Isaacs general model, or even significant variations. Salley's (1977) review of previous studies reflects considerable difference in procedures but not in concept or basic approach. There have been, however, some interesting special studies conducted. Three are mentioned here. A Texas institution attempted to estimate the local economic impact that would result from various funding level changes. West Virginia University studied the impact of non-resident student spending and then compared it with the impact of the local tourist industry. The University of Illinois-Urbana studied the impact of student Christmas buying, the results of which had implications for the school calendar.

Basically, the Caffrey-Isaacs model attempts to account for the total cash flow into and out of the institution and to calculate the net direct and indirect economic impact on some predetermined geographical area, usually a local service district. Sub-models are available to calculate a number of different kinds of impacts, including those affecting local businesses, government, and individuals. These result in 20 different measures of impact on the local economy. The most frequently used sub-models are those related to impact on the business community and the most common measures, additional income and jobs. The model, or variations of it, most frequently used is shown in Figure 1.

Figure 1

Basic Elements in Determining Economic Impact



Some key terms are the following:

- Direct Impact: actual dollars spent locally by the college, students and faculty/staff.
- Indirect Impact: an estimate of the amount respent, or turned over associated with the original purchase or expenditures.
- Total Economic Impact: direct plus indirect impact.
- Multiplier: a ratio which estimates the extent to which initial expenditures by the institution, staff and students are respent by businesses and individuals within the local economy. See Salley (1977) and Posey (1986) for a thorough description of procedures for calculating the multiplier.
- Economic Return: usually expressed as a ratio of the funds allocated to the college or system to the total economic impact; for instance, if the rate of return is 3 to 1, this means that for every dollar invested in the college, three are returned to the local or state economy.

Data Sources and Requirements

These can be fairly limited or staggering depending on how much of the basic model is implemented. Local resources and capabilities must be reviewed in light of data requirements associated with different parts of the basic model. For the limited model implemented by most studies, the major data sources are college records, including a record of purchases, and student and faculty/staff surveys, used to derive estimates of local spending. The surveys can be very straightforward or become more complicated if it is decided, for instance, to derive estimates of faculty/staff savings in local financial institutions. In addition, and in this example, invasion of privacy issues may be raised. The Caffrey-Isaacs reference identifies the data requirements of the various sub-models.

For the most part, survey design issues and those associated with deriving reliable estimates of local college spending require adherence to sound survey research practice, including appropriate sampling strategies. Some studies rely exclusively on available records and exclude surveys to estimate spending patterns. Many of these studies, while less costly to conduct, are also less reliable and may be viewed as less credible by those receiving the reports.

The Multiplier

Carol Channing in "Hello Dolly" said: Money is like manure. It's no good unless you spread it around.

The "multiplier" is a number that helps to estimate how much initial spending is "spread around" locally. More formally put, the multiplier indicates how much respending of local college, staff, and student expenditures takes place. How the multiplier should be calculated continues to be a thorny issue.

Many studies use a multiplier somewhere in the range originally suggested by Caffrey and Isaacs, 1.2 - 1.5 (this figure is multiplied by the direct impact to determine total economic impact). Others simply use a number used previously by a college or system similar to their own or select a number somewhere in the range of numbers used in a variety of similar settings. Generally, the larger the community, the more diverse the employment distribution, and the less the area is dependent on imports, the higher the value of the multiplier.

Salley's (1977) important work provides a specific method and formula for deriving the multiplier using payroll/local sales ratios, available from institutional records and census data. Sometimes, another agency in the local area or state has developed an estimate of the respending rate that may prove useful to study directors.

SOME ADDITIONAL ISSUES

When moving from estimating impacts on the local to the statewide economy, the "export sector" concept underlying the Caffrey-Isaacs model becomes strained. For example, only students from outside the local area can be included in the model. Thus, when attempting to describe statewide impact of a system or a group of local community colleges, only students from out-of-state, or those who would elect to attend an out-of-state college were the institution not there can properly be included in the model. Also, different multipliers are needed for local versus statewide studies since dollars will leave the local region faster than they will leave the state, a more diverse and larger sector of the economy. Most previous studies are conducted at the local level, thus minimizing some problems.

The specific procedures used and what is selected for emphasis has varied by type of institution. Institutions with large numbers of students from out-of-district and out-of-state have paid attention to the impacts associated with local spending by parents, alumni, and tourists attracted to the area because of the institution. Community colleges have shown particular interest in the impact on employment. While the model is most strained with community colleges and commuter four-year urban institutions, these colleges have been particularly active conducting these studies.

SELECT PROCEDURAL AND TECHNICAL ISSUES

There are eight specific issues that can influence the validity, reliability, and credibility of a study.

1. Surveying Students and Staff

The key issue here is estimating college related expenditures. What constitutes "college-related"? Is it enough that a student perceives a particular expense to be college-related? Little is known about the reliability of student estimates of spending, especially in determining the amount that is college-related. Most studies using the export sector model simply collect student spending data and apply the results to the proportion of out-of-area students. Students are not asked to estimate the amount of their spending that is "college-related." Salley (1977) references one study that used student diaries to determine expenditure patterns.

In Oregon the community colleges are attempting currently to include housing costs only when students indicate these costs are directly associated with attending the community college. If these costs would be incurred if they were not attending the college, the costs are to be excluded.

Most studies require students and sometimes faculty/staff to breakout spending by type (for example, housing, food, etc.). This is done for two reasons. First, such breakouts are needed for part of the full model. Second, the breakouts are used as a means of compiling more accurate information from the respondents.

Four-year colleges and universities need to weight student and staff expenditure sample data by type, freshman versus graduate student, and full-time versus part-time staff, since spending patterns will most likely differ among these groups. Thus, a weighted overall average is used in subsequent calculations.

2. Use of an Outside Consultant

To enhance the credibility of the study, and to minimize the perception that it is self-serving and not reliable, institutions have used several strategies. Some have recruited community members to serve on the local study committee. Some have employed an outside consultant to audit the study process and results. Others have recruited the expertise of a local economist to assist with multiplier derivation. And still others have simply contracted the entire study to an outside firm. There is probably no one best way to enhance credibility, but one of these alternatives should be considered.

3. The Multiplier

Quite a bit has already been said about this issue. Salley's (1977) work stands as the most comprehensive analysis of and prescription for calculating the multiplier. The final test of the multiplier to be used is one of reasonableness. Does it fit with previous studies? Does it fit with any local studies conducted to derive a multiplier estimate? Very small changes in the multiplier

can greatly change the final dollar or job impact figures. Thus, when in doubt, err toward the conservative.

4. Part-Time Students

The question here is, of course, should part-time student expenditures be included in the model. Some part-time student spending is undoubtedly college-related. Can it be assumed, however, that these expenditures would leave the local economy if not for the existence of the college? Would money not spent on college related activities be saved for or spent on other activities within the local economy? Part-time students should probably not be included unless they are known to have relocated to the area for the purpose of attending the college or, assuming that the college did not exist, they would leave the local area.

5. Sources of Double Counting in Deriving Direct Spending Figures

A helpful first step in laying out the study is to draw a schematic depicting exactly what is being accounted for where in the model. Most particularly, the categories of student expenditures (for example, tuition and fees, housing, books, etc.) must be compared with college expenditures to avoid double counting in areas such as bookstore, housing, other auxiliary services, financial aid, etc.

6. Boundary Definition

Another important first step in any study is defining geographic boundaries. This is not a problem for most community colleges or for colleges located in small towns. It is a problem, though, for colleges and universities located in the center or the periphery of large metropolitan areas. Most use either the city core for its boundary or the larger metropolitan area, depending on purposes and audiences for the study. And others find merit in also calculating statewide impact. Whatever the definition, it must be precise and communicated unambiguously to those completing surveys that ask whether the respondent lives in or out of the "district." Many surveys include a map on the back of or attached to the questionnaire.

7. Coordination

Coordination of a study can be particularly challenging, especially when local data is to be collected separately and then cumulated in a systemwide or statewide study. The Oregon community college experience (1982) offers some advice on procedures, including the appointment of a study director, campus-based coordinators and a statewide advisory committee consisting of the director and local coordinators, and the careful development of written procedural guidelines.

At the institution level coordination can also prove challenging. The study effort requires three critical ingredients: technical expertise; support from students, faculty and staff, especially during the period of time when surveys are to be administered; and involvement by those who will disseminate the results of the study. A broad-based local advisory committee should include those representing each of these groups. Inclusion of a business office representative and a public information officer is especially important.

8. Negative Impacts

Most studies do not examine negative economic impacts and probably should not unless the full Caffrey-Isaacs model is implemented. To do so could result in distortions, largely because most studies do not take into account all of the impacts on the positive side. Examples of negative impacts are property and sales tax foregone due to the tax-exempt status of the institution, the costs of municipal services to the institutions and increased tax burdens attributable to the services which must be provided to college employees, students and their families (Erwin & Miller, 1982, pp. 20-21).

HOW THE RESULTS HAVE BEEN USED

The predominant use of results has been for public relations purposes. Some of the ways study results have been disseminated include: press releases; special brochures aimed at a particular audience; and full reports, in booklet or pamphlet form, that describe the institution's or system's contribution to economic development and/or community development.

The use of graphics has been especially effective in communicating results of the study. One college in Oregon, for instance, shows a coin with a pie-like slice set apart representing the state's and district's investment and the remainder of the coin representing what the college gives back to the local economy. The headline of this brochure developed by Lane Community College was "invest a dollar, get back four," the kind of message highlighted by many institutions completing these studies (see Kinnick, 1982, for further examples from Oregon community colleges.). Many studies report results in ratio form, as a dollar amount representing a return on a $1.00 investment by state and local taxpayers.

How the study results are interpreted is especially critical. For example, unless the full export sector model is used for the study, it is not accurate to say that all of the economic impact would be lost if the college or university did not exist. In the case of commuter institutions, a portion would stay in the local area.

Special care must be taken to point out that while important, the economic impact of the institution is not the

reason for its existence. Enthusiasts, in the rush to call attention to dollar and job impact on the local area, may forget to remind the reader or listener that the primary mission is educational.

RESOURCES

Institutions or systems similar to your own that have conducted economic impact studies are relatively easy to locate through the ERIC system. The following references include sources that are especially helpful in terms of conceptualizing the study and developing procedures.

1. Concepts, Models

Alfred, R. (1982). Institutional impacts on campus, community and business constituencies. San Francisco: Jossey-Bass.

Bowen, H. (1977). Investment in learning. San Francisco: Jossey-Bass.

Caffrey, J. & Isaacs, H. (1971). Estimating the impact of a college or univeristy on the local economy. Washington, DC: American Council on Education.

National Center for Higher Education Management Systems (1978). Conducting community impact studies. A Handbook for Community Colleges Boulder: NCHEMS.

Salley, C. D. (1977). Calculating the economic multiplier for local university spending. In R. J. Fenske, & P. J. Staskey, (Issue Eds.), Research and Planning for Higher Education. Tallahassee, FL: The Association for Institutional Research.

SRI International, Public Policy Center (1986). The higher education-economic development connection: Emerging roles for colleges and universities in a changing economy. Washington, DC: American Association of State Colleges and Land Grant Universities.

2. Bibliographies, Guidebooks, Special Articles

Booth, G. & Jarrett, J. (1976). The identification and estimation of a university's economic impacts., Journal of Higher Education, 47, pp. 565-576.

Erwin, J. M., & Miller, J. L., Jr. (1982, May). An analysis of state-level studies of the economic impact of higher education. Paper presented at the 22nd Annual Forum of the Association for Institutional Research, Denver, CO.

Fink, I. S. (1976). The economic impact of institutions of higher education on local communities: An annotated bibliography. Berkeley, CA: University of California. ED138128

Kinnick, M. (Ed.) (1982). Oregon Community College Economic Impact Study: A Guidebook. Gresham, OR: Mt. Hood Community College.

Posey, E. (1986, June). Conducting an economic impact study. Materials used in a workshop conducted for the 26th Annual Forum of the Association for Institutional Research, Orlando, FL. (Available from the Office of Institutional Research, Georgia State University, Atlanta, GA.)

Ryan, G. J. (1985). A shortcut to estimating economic impact. Community and Junior College Quarterly of Research and Practice, 9, pp. 197-214.

INSTITUTIONAL RESEARCH SUPPORT OF THE SELF STUDY

Richard D. Howard
North Carolina State University

James O. Nichols
University of Mississippi

Larry W. Gracie
North Carolina State University

INTRODUCTION

The role of institutional research in the self-study process varies from institution to institution. The role of any specific office will vary from supplying data to the self-study committee to that of coordinating and directing the self-study (Peterson & Corcoran, 1985). Obviously, how a single office fits into the self-study process is a function of its role in day-to-day decisionmaking and planning activities. Institutional research can support the self-study process in accreditation activities in numerous ways. Before getting into the specifics of the self-study, a review of the general purposes and processes of accrediting activities is presented. This is then followed by a suggested outline for conducting a self-study.

ACCREDITATION - WHAT IS IT?

Voluntary accreditation is unique to American higher education (Christal & Jones, 1985). It provides a means for self regulation and has evolved as a major force for enhancing academic and educational quality. There are two types of accrediting agencies - institutional and professional. Institutional accrediting agencies are concerned with the institution as a whole. These are typically regional agencies and are directed by member institutional representatives. As such, review of any given institution by one of these agencies is usually conducted by academic and/or administrative staff of various regional institutions.

Professional accreditation on the other hand deals with the review of specific programs. In these cases visiting teams are often staffed by both academic and non-academic personnel. In most cases, professional accreditation is contingent on institutional accreditation. For ease of presentation, the rest of the chapter will speak to the institutional accreditation self-study support. In general, however, the activities described are also appropriate for professional accreditation.

Until recently, accreditation was based primarily on quantifiable input information about the institution as it related to the primary mission of the institution. In this approach, the processes and resources available for

meeting the institutional mission were studied. Recently, the focus of accreditation activities has begun to shift to the outcomes of the educational process and programs. As this component becomes a part of the accreditation process, it will become the mandate of the institution's self-study to examine not only the adequacy of resources to support the institution's programs and processes but also the quality of the institution's outcomes. (This has become a reality for institutions in the Southeast for those institutions accredited by the Southern Association for Colleges and Schools [SACS].) It is through the self-study that evaluations of both input and outcomes will be conducted.

The rest of this chapter is concerned with a proposed self-study process. While no single component is more or less important than any other, more emphasis has been placed on those parts of the study which fall within typical institutional research functions (see Saupe). Underlying this discussion is the notion that typical research, planning, decision support, and reporting activities of institutional research will provide a college or university with much of the base line data that will be required in a self-study. As such, as an institution begins its self-study, it may prove useful to review projects and reports that have been developed in the past four to five years in order to determine potential research needs and sources of data.

SELF-STUDY

A primary result of the accreditation self-study process should be the identification of the institution's strengths and weaknesses. This is both a qualitative and quantitative examination and evaluation of the institution's progress toward its mission and meeting its goals and objectives. To accomplish this, the process of a self-study can be conceptualized in much the same manner as the typical research or evaluation project.

RESEARCH/EVALUATION DESIGN	SELF-STUDY DESIGN
Statement of the Problem	Define/Reaffirm Mission
Literature Review	Area Reviews by Subcommittees
Development of Research Hypotheses	Development/Review of Area Goals and Objectives
Methodology:	Methodology:
Instruments	Instruments
Design	Design
Sample	Sample
Procedure	Procedures
Data Analysis	Data Analysis
Review of Results	Review of Results
Recommendations	Recommendations to Strengthen Identified Weaknesses
Implications for Further Research	Develop Evaluation of Recommendations

Using this analogy, the first step in a self-study would be to develop the framework within which the study is to be conducted. This qualitative step must be accomplished by senior administrators and/or the governing constituency of the institution. At this stage, the institution's mission should be reaffirmed or developed. This is analogous to developing the research questions in a research project and will provide the context in which the self-study is conducted

Once the self-study committee has developed or reaffirmed the mission statement of the institution, a number of subcommittees will be formed to examine specific operational components of the institution (academic programs, student affairs, etc.). The mandate of these subcommittees is to review the goals and objectives specific to the areas of their concern and to identify appropriate indicators to measure the quality and quantity of each area's outcomes. Institutional research at this point can help develop strategies and identify the variables to measure progress toward the stated goals and objectives.

DATA SOURCES

Once the organizational framework has been developed for the self-study, specific data elements have to be identified. Data to support the self-study can come in many different forms and through different types of media. More and more the computer is becoming the tool through which most data are stored, manipulated, and analyzed. As such, when data base development in preparation for a self-study is discussed, computerized files are usually the first resource considered. Remember, however, that at most institutions there is a wealth of data/information that is not computerized. In addition, data sources external to your institution are often transmitted on paper (although many national and international data bases are being made available in computerized form). Below are several types of data commonly used in both the management and planning of the institution that can be reexamined for the specific purpose of the self-study. At this point, it should be recognized that institutional research activities have been, for the most part, a function of quantitative data collection and analysis. Therefore, most of the sources of data discussed below are quantitative. This is not to suggest that qualitative data is not an important part of an institution's self study.

Operational and Census Files

Operational and census files are the computerized files which are typically maintained and created by the operational units within your institution (admissions, registrar, personnel, etc.). The operational files are dynamic in that they are updated everyday. As such, if one accesses data from these files over a period of several days, different results may be incurred in each analysis. Census files, on the other hand, are snapshots of the op-

erating files. The purpose of the census file is to provide a consistent source of data from which various analyses can be conducted throughout the year. The other strength of the census file is that, if created on the same day from semester-to-semester or year-to-year, trend data can be developed over a period of time. The trends can then be analyzed without having to worry about variations due to the artifacts of various processing calendars from one year to the next. It is recommended that reporting, planning, and decisionmaking at senior levels be supported by data and information developed from census files to assure consistency between reports and to lend credibility to both the analyses and policies developed from them.

Existing Reports

During the past twenty years, higher education institutions have been required to report a great deal of data to various federal, state, and system agencies. Every institution has had to submit HEGIS/IPEDS reports to the federal government. Because of the consistency in data element definitions of these reports, these same data can be a valuable historical resource for student, personnel, and financial data. These data tend to be even more creditable because they were put together for federal reporting requirements and use consistent data element definitions each year. In addition, there should be a number of reports developed for the governing constituencies of each institution. A third type of report is past accreditation reports whether they be regional or professional. Again, if these reports are or have been developed using census file data, their consistency will help to provide a creditable source of information to the self-study. Review of these or similar data/studies provide insights to the history of the institution and establish a base for projecting the future.

Surveys

Surveys, in general, fall under two categories. The first is a one-time survey which tends to address a specific question or concern. Over a period of years, faculty, staff, students, employers, and others may have been surveyed in response to specific internal or external concerns of the university. These studies, while developed to answer a specific question, may provide data relevant to accreditation demands. It is almost certain that special surveys will be requested by one or more of the various subcommittees.

Periodic surveys are the second category and tend to be done in a systematic sequence. These surveys are usually interrelated. Their intent is to provide a picture of a certain component of the institution either across time or across constituencies. An example of this would be a series of graduate surveys. On a yearly basis, one may survey the graduates as they are leaving the institution. Follow-up surveys of graduates who have been out for three

years would be conducted on a three-year cycle and on a five-year cycle to those who had been out for five years. In each case, asking the same or similar questions may provide valuable insights about how former students value their experiences at the institution and if they change over time. In addition, the analysis can include comparisons between graduates from different disciplines. The use of surveys can provide a great deal of information to the self-study team in assessing the impact of the institution's programs. Survey research, while one of the more common forms of research in the social sciences, is also one of the more tricky. A word of caution: spend a good deal of time both in planning surveys and in developing the survey instruments.

Make sure that your questions will provide information which will contribute to assessing the institution's progress toward meeting its stated goals and objectives. Several references are provided at the end of the chapter which deal specifically with survey research methodology and data analysis (Babbie, 1973).

Past survey efforts, in addition to those done in direct response to the self-study committee or a subcommittee, often will provide the most critical information for evaluating outcomes described in the various goals and objectives and the mission statement. It should be noted that with the recent increased interest in "outcome measure" more institutions are implementing plans for ongoing survey projects to evaluate the institutions goals and objectives (Ewell, 1985; Miller, 1981).

Environmental Scanning

Environmental scanning has received a good deal of attention in the past several years. In essence, the objective is to assess the environment in which your institution exists (Callan, 1986). Data typically collected in environmental scanning deal with the potential student pool, financial conditions, political pressures, potential industrial growth, or any other condition or pressure which may impinge on the offerings and operation of the institution. While a good deal has been written about how to scan the environment, it seems that functional models have yet to be developed which allow the results of environmental scanning to be quantitatively incorporated into the university's data bases and made a part of decision/planning support systems. In many instances, the analysis of environmental data is a qualitative process as senior officials review the data and through discussions evaluate its potential impact on the institution.

Peer Data

One way in which to evaluate the condition of your institution is to compare it to other similar institutions. Data about institutions or higher education in general can be found or collected from a number of sources. It is

beyond the scope of this chapter to provide a comprehensive list of these sources; however, several examples are presented which are typical of the comparative data which are available. Virtually all comparative data are developed as a result of organized data exchanges between institutions or the result of regional or federal reporting mandates. In the case of data exchanges, the comparative data are typically made available only to the institutions which have participated in the exchange of data. In many cases, the data exchange processes have been developed by a group of similar institutions to provide specific types of comparative data. National or regional Association for Institutional Research (AIR) meetings have often been the forum through which the participants have met and developed their processes for data exchanges. (This has been and continues to be a very important outcome of professional meetings.)

A second source of comparative data is through regional and national data collections. The best known is the HEGIS/IPEDS data. These data, along with AAUP salary data, are published in several forms - the most popular being The Chronicle of Higher Education. These data bases can also be requested from the Center for Statistics of the U. S. Department of Education. There have been questions raised about the consistency of the HEGIS data between institutions. The Center for Statistics has tried to address some of these concerns with the IPEDS data collection. (This replaced the HEGIS data collections.) The 1984 Firnberg and Christal analysis of HEGIS data bases identified reliability and validity problems that should be considered when using the data. There are a number of such data collections each year with the data being made available in a number of formats to interested parties. A publication by the State Higher Education Executive Officers (SHEEO) entitled Comparative Data About Higher Education: A Resource Directory, is an excellent source of data bases which are available. (Also see Christal & Wittstruck, 1986; Halstead, 1979.)

Often, a serious problem will be to identify an appropriate group of peers. Comparative analyses will do little to help evaluate the strengths of the institutions or programs if senior decisionmakers do not agree with the institutions used in the comparison. Special care needs to be taken in the selection of peers. Quantitative algorithms have been developed by NCHEMS and the Institutional Research Office at the University of Kansas to select peers. In both cases, HEGIS data are used as the base for the selection (Teeter & Christal, 1985). Regardless of how the comparative institutions are chosen, it is imperative that there is agreement from senior officials on which peer institutions are to be examined.

A Note About Qualitative Data

The environment in which decisions are made and the impact of data used in these discussions are more and more be-

coming a critical consideration of the institutional review. Examination of "who makes key decisions" and "how decisions are reached" is a way to assess the validity of research questions. Therefore, review of the models used to reach decisions on critical matters has become important. Data sources which may assist in analyzing these processes include minutes of special task force meetings, studies that were conducted in support of specific decisions, plans to evaluate outcomes, documentation of how costs are determined, what alternative plans were considered, and how actual outcomes will be evaluated in respect to projected outcomes. New qualitative research tools, such as content analysis, may need to be developed in the institutional research office.

ANALYSIS AND PRESENTATION OF DATA

Once the data have been collected, the next step is to analyze and present the findings in some fashion to the subcommittees, the self-study committee, and perhaps for the actual self-study report. As the purpose of this chapter is to give an overview of how institutional research can support the self-study procedure, there will be no detail about various forms of analysis or presentation modes. The bibliography suggests a number of references to help with both the analysis as well as presentation of the data. The analysis and results should be presented in the simplest format that will quickly convey the message you are trying to get across. Presentations of means, standard deviations, and frequencies will typically satisfy most members of the self-study committee. This is not to say that more indepth analysis should not be done. Also, data analyses are no longer solely reported in tabular formation - graphic presentations have added new dimensions to data presentation (Tufte, 1983).

One of the most popular forms of presenting data is that of a factbook. A factbook developed yearly will, over a period of years, provide a historical context from which trends can be identified. Factbooks take many forms and are used in many different ways. Look to your institutional research colleagues for suggestions on how a factbook may be developed. In addition, do not hesitate to ask for a copy (see Chapter 9).

AN ADDITIONAL RESOURCE

The final, but perhaps the most useful, resource is that of your institutional research colleagues. Membership in either the national, regional, or state institutional research associations will put you in touch with many institutional research professionals. Other professional associations such as the American Educational Research Association (AERA), the Society for College and University Planning (SCUP), and CAUSE often will provide resources that can support self-study projects. Use these colleagues as a resource. A quick telephone call or a short letter is typically all that is required to get comparative

data or answers to questions that you may have. For best results, try to request data that can be found in existing reports, that is, HEGIS/IPEDS reports, AAUP reports, or other standard reports. If possible, a trip to a similar institution that has recently completed its self-study will prove valuable. Be assured that virtually any question or situation has most likely been encountered somewhere else.

EVALUATION OF THE SELF-STUDY RESULTS

This is the final stage of the self-study and is analogous to the conclusions and recommendations found in most research reports. As in the first stage of the self-study, this qualitative analysis is typically done by senior officials of the institution. The role that the office of institutional research will play in this particular step of the study will vary depending upon its overall role in the self-study process as discussed above. Two particular concerns must be addressed in this final stage of the self-study. The first is making sure that the minimal data requirements, defined by the accrediting agency, have been collected, tabulated, and summarized in the appropriate format. For the most part, it is a straightforward process once the data are collected. When the specific data requirements have been met, their analysis and qualitative review in conjunction with other data make up the second component of this final stage of the self-study.

The second and perhaps more important component of this stage of the self-study is the interpretation of the results of the study relative to the goals and objectives of the institution. This "evaluation" of the institution's outcomes should result in the identification of both strengths and weaknesses. It is critical that weaknesses be studied and their causes be identified. At this point, it is incumbent upon the institution to identify corrective actions and document their incorporation into normal processes. It is very important that this action be taken by the institution before the visiting team arrives on campus. While institutional research typically will have very little input into this facet of the evaluation, it is important that those responsible for the institutional research function be aware of decisions and actions taken as a result of the self-study. They may be called upon later to evaluate the effects of the actions.

An institution may very well meet the various minimum requirements or standards of any given accrediting agency and still not be accredited; however, once these minimal standards are met, it is the evaluation and subsequent corrective actions of the institution which will establish the degree of success the institution experiences in the accreditation process.

CONCLUSIONS

More and more, traditional institutional research activities are becoming important support components in self-studies and the final accreditation decision. Studies to ascertain the quality and success of the institution as defined in their goals and objectives are being examined by the accrediting agencies. While the role of specific institutional research offices may vary from institution to institution in the self study process, it is inevitable that they will be drawn into the process. This chapter has illustrated how normal activities of an institutional research office can provide much of the data/information needed to support a self-study. A number of references are listed at the end of this chapter. Some are technical in nature and refer to the analytical support necessary to conduct surveys or develop trend analyses. Others will provide an overview of accreditation processes in America. Remember, however, that the best resource is often the colleague met at the last professional meeting.

It should be apparent that the self-study process described above is, in essence, the process one would use in the conduct of a summative evaluation. This is substantially different from earlier accreditation research requirements which resulted in a formative or process evaluation. Increasingly, institutions are being required to assess the quality of their programs and their outcomes through continuous reexamination. It is appropriate for institutional research to provide an objective presence in both the design of various components of the self-study and the evaluation of findings.

REFERENCES

Babbie, E. R. (1973). Survey research methods. Belmont, CA: Wadsworth Publishing Company, Inc.

Callan, P. M. (Ed.). (1986). Environmental scanning for strategic leadership. New directions for institutional research, 52. San Francisco: Jossey-Bass.

Christal, M. E. & Jones, D. P. (1985). A common language for postsecondary accreditation: Categories and definitions for data collection. Boulder, CO: National Center for Higher Education Management Systems.

Christal, M. E. & Wittstruck, J. R. (1986). Comparative data about higher education: A resource directory. Denver, CO: State Higher Education Executive Officers.

Ewell, P. T. (Issue Ed.). (1985). Assessing educational outcomes. New directions for institutional research, 47. San Francisco: Jossey-Bass.

Firnberg, J. W., & Christal, M. E. (1984). HEGIS data: Is what you see what you get? College and University. 60(1).

Halstead, D. K. (Ed.). (1979). Higher education planning: A bibliographic handbook. Washington, DC: US Government Printing Office.

Kells, H. R. (1980). Self study processes: A guide for postsecondary institutions. Washington, DC: American Council on Education.

Miller, R. I. (Issue Ed.), (1981). Institutional assessment for self-improvement. New directions for institutional research, 29. San Francisco: Jossey-Bass.

Peterson, M. W., & Corcoran, M. (Issue Eds.), (1985). Institutional research in transition. New directions for institutional research, 46. San Francisco: Jossey-Bass.

Saupe, J. L. (1981). The functions of institutional research. Tallahassee, FL: The Association of Institutional Research.

Teeter, D. J., & Christal, M. E. (1985). Establishment of peer groups: Is there a best method? Paper presented at the Association for Institutional Research Forum. Portland, OR.

Tufte, E. R. (1983). The visual display of quantitative information. Cheshire, CT: Graphic Press.

PEER INSTITUTIONAL STUDIES/INSTITUTIONAL COMPARISONS

Deborah J. Teeter
University of Kansas

Paul T. Brinkman
National Center for Higher Education Management Systems (NCHEMS)

The exchange and comparison of institutional data have become more pervasive on the higher education scene. The information explosion and information-oriented management are in part responsible for this growth. There are a number of issues for which comparative data are desirable such as resource utilization patterns (student-faculty ratios, class size); input patterns (faculty-support staff ratios); output patterns (degrees awarded, research expenditures per faculty, attrition/graduation rates of students); and relative efficiency measures (expenditures per credit hour). Data from other institutions provide management with the ability to size up competition, benchmarks for assessing the well-being of their own institution, the ability to pinpoint areas deserving attention, and guides for policy development. To satisfy increased demands for accountability by local, state, and federal agencies, institutions can use comparative data to couch responses in the proper perspective. Comparative data are being used with increasing frequency to explain and justify budget requests, salary increases, teaching loads, and tuition increases.

The motivations for exchanging and comparing data are keys to the specifics of data comparisons and the selection of institutions with which to compare. This brief introduction has alluded to some of the reasons for comparing data. The next three sections of this chapter will describe the nature of data comparisons, types of comparison groups, and procedures for selecting comparison institutions.

ASPECTS OF COMPARATIVE DATA

Before discussing the nature and type of data compared, it is useful to distinguish between comparative data and comparative information. Data may take the form of numbers that represent measurements of continuous quantities. Or data may consist of codes that represent observations of the discrete presence or absence of certain characteristics. (Enrollment data take the form of numbers; a student's religious preference is recorded by a code as would the presence of a process or policy.) Information consists of data that are useful in one or more contexts, that inform someone about something, that reduce uncertainty in some way. To become information, data may have to be combined in one or more ways, analyzed, synthesized, or displayed in a particular form, although this is not always the case. To determine whether or not a planned level of overall institutional full-time equivalent (FTE)

enrollment has been reached, only one number, a simple total, is needed. By contrast, to know something about the pattern of increases and decreases in enrollment in the various student-major programs offered in colleges of arts and sciences, the total enrollment figure must be broken down to department-level numbers. Then each of those numbers must be compared to its counterpart number for the previous year (Brinkman & Krakower, 1983).

The focus of the discussion that follows is on data as the basic building blocks for relatively formal comparisons in higher education. A more complete discussion of helpful distinctions between data and information can be found in Jones (1982).

Types of Data for Comparisons

The type of data most frequently exchanged among institutions are facts and figures about a particular topic. Enrollment, appropriations and other financial data, faculty salaries, tenure policies, and general policies and procedures are often of interest to other institutions. For example, enrollment trends in a particular discipline might be of great interest to an institution that is experiencing declines it believes to be counter to national or regional trends. Comparative enrollment data help determine whether the problem is local or of a more general nature. Information about the procedures and policies instituted by institutions that have had to deal with financial reversals might be helpful to institutions which have yet to experience the same kinds of financial crises. The generic form of the questions being asked is "How well are we doing as compared to institution X?"

The kinds of data frequently exchanged about faculty include not only salary information but also tenure percentages. Student-faculty ratios and faculty activity analyses are examples where comparative data substitute for well-defined norms for faculty workload and productivity. However, while these ratios and rates can readily be calculated, it is frequently difficult to interpret them unambiguously. Often the reasons for the ratios can only be guessed at, as they typically reflect subtle combinations of historical accident, funding levels, management decisions, program mix, and so on. However, information from similar institutions can indicate whether these ratios and rates are at or near group averages or other norms (if available) and, if they are not, whether additional investigation is warranted.

Since institutions throughout the country are frequently subjected to the same governmental regulations, particularly if they have similar missions, it might be useful to exchange approaches to meeting these requirements. For example, several years ago when institutions needed to develop procedures to meet the requirement of the Office of Management and Budget's Circular A-21 regarding indirect cost recovery on federal grants, informal polls were con-

ducted about how institutions were responding to the A-21 requirements. Institutions that had decided how to meet the requirements were often willing to share with others details of their proposed procedures.

Research findings and analysis on a particular topic might also be exchanged. Institutions concerned with attracting new students and retaining current students might be able to borrow research findings from others or implement recommendations without redundantly conducting the same research. Special research projects can be both time consuming and costly; often an institution may not have the time or resources to conduct its own research. While borrowing other institutions' data may be somewhat risky, careful examination and application of research findings could save time and money. A prior comparative analysis might reveal whether the institutional setting within which the research was conducted is sufficiently similar to suggest that research findings are likely to be transferable. Such comparative analysis is, of course, no defense against a bad piece of research (Teeter, 1983).

Gathering Data

There are several ways in which comparative data can be acquired: 1) use published data such as those collected annually by the Center for Education Statistics (see Christal & Wittstruck, 1987, for a list of data sources); 2) collect data on an ad hoc basis from selected institutions in order to respond to a particular problem or issue in a timely manner (in this case, all the benefits accrue to the institution requesting the data); and 3) exchange data for the mutual benefit of participating parties.

The formal exchange of data typically makes possible the sharing of more complicated data. Joint development of data formats, data definitions, and exchange procedures are not uncommon in efforts to enhance the comparability of the shared data. Data exchange networks avoid redundant requests for data among institutions that are interested in the same type of comparative data on a routine basis. Routine exchange provides a baseline of data for trend analysis and often incorporates explicit guidelines for the use and sharing of the data. These guidelines can promote cooperation, because institutions appreciate having some assurance of how their data will be handled.

Limitation of Data Comparisons

Fundamental data concerns having to do with validity, accuracy, and reliability are present in the comparative context as in any other situation in which data are to be taken seriously. Establishing how well these concerns are met is often more challenging when doing comparative analysis: comparative data are often derived from multiple sources; the rules and definitions for recording such data may be inconsistent across sources; and the close famili-

arity that can be so helpful in spotting data errors is usually missing because one typically must depend on secondary sources. The use and purpose of the comparison determines in part the extent to which errors of a given kind may compromise the comparison. For example, management-control situations may require highly accurate data, while data that are to be used in a strategic-planning context could probably be less accurate, in the sense of being precise, without causing problems (Brinkman & Krakower, 1983).

Politics of Using Comparative Data

Since comparative data are often used to justify, explain, or advocate a certain position, it is important to understand how the intended audience, whether internal or external to the institution, looks upon comparative data. Organizational and political realities need to be considered and a strategy developed accordingly. If the audience is likely to be hostile to the idea of using comparative data, it is imperative to involve them early in the game. Concerns often center on the validity of the comparison group. If that is likely to be the case, be sure to include the audience in the selection of comparison institutions. It might serve to mitigate some of their concerns if they understand the rationale and criteria used to select the comparison institutions. Obviously, data are expected to be accurate and sufficiently detailed for the audience to properly assess the implications of the data.

The key to successful use of comparative data is properly sizing up the environment in which the data are to be used and taking the steps necessary to insure that the audience will be receptive. Failure to lay the proper groundwork may cause extensive delay in the use of comparative data or prohibit their use altogether (Brinkman & Teeter, 1987).

TYPES OF COMPARISON GROUPS

There are a number of different types of comparison groups and all of them can play a legitimate role depending upon the situation. The issue to be addressed by comparative data and the circumstances within which such data are to be used will be important considerations in choosing the type of comparison group. Numerous typologies for comparison groups can be envisioned. The discussion below distinguishes four types of groups - competitor, peer, aspirational, and predetermined - from Brinkman and Teeter (1987).

A competitor group consists of institutions that compete with one another for students or faculty or financial resources. Competitors may not necessarily be similar in role and scope, although they frequently are. Depending upon the purpose of the comparison, their lack of similarity, if that should be the case, may or may not be important. Comparisons that depend upon institutional

similarity could be at risk if a competitor group is used without further investigation.

A peer group consists of institutions that are similar in role and scope, or mission. In this case, "similar" rather than "identical" is the operative word. It is unrealistic to expect to find clones of the "home" institution, but perfectly reasonable to expect to find other institutions that match sufficiently on particular characteristics. Sufficiency in this context is ultimately a matter of judgment.

Comparison, by definition, means examining both similarities and differences. The latter is critical in developing an aspiration group. The term refers to a group that includes institutions that are dissimilar to the home institution but worthy of emulation. When a comparison group contains numerous institutions that are clearly superior to the home institution, the group reflects aspiration more than commonality of mission.

One word of caution in using aspiration groups: if they are presented as if they were peer groups, they can be costly in the political arena outside the campus. For example, if the use of comparative data to buttress resource arguments is an issue, the masquerading of an aspiration group as a peer group will put at risk the credibility of most any comparative data the home institution wishes to use. It may fall to the institutional researcher to provide the reality check through assembling and presenting objective data on the alleged peers. If the appropriate data are chosen, the aspirational character of the proposed comparison group will usually be obvious. Still, individuals will ignore the obvious under certain circumstances, such as a no-holds-barred effort to increase funding, so the institutional researcher is advised to size up carefully where the comparative analysis fits in the overall institutional strategy. It cannot be stressed too much that the task of moving comparative data into the decisionmaking process is often more important than the technical routines for working with comparative data.

Predetermined institutional comparison groups can be thought of as falling into one of four categories: natural, traditional, jurisdictional, and classification-based. Natural groups are those that are based on one or more of the following types of relationships: membership in an athletic conference; membership in a regional compact; or location in a region of the country. In effect, institutions already belong to a highly visible grouping of some sort, so it is natural to think of them as being comparable. They may indeed be comparable in some inherent sense, but the specific nature of the comparison is the critical test as to whether a natural comparison group will suffice.

A traditional comparison group is one that is founded in history - its only virtue is that it has been used for a

long time. It may or may not be an appropriate comparison group in a given situation.

A jurisdictional group consists of institutions that are compared simply because they are a part of the same political or legal jurisdiction. Frequently, the boundary for this type of group is the state line. Not surprisingly, elected officials and state agency staff will make comparisons of institutions within their purview, even though the institutions may have little else in common. Once again, the comparison issue in question is (or should be) the primary factor in determining the appropriateness of this kind of comparison group.

Institutional classifications used for national reporting can also be the basis for choosing comparison institutions. Probably the best known is the classification developed by the Carnegie Commission in the 1970s (Carnegie Commission on Higher Education, 1976) and currently being updated. The American Association of University Professors uses an institutional classification for reporting comparative faculty salaries (see the annual issue of their journal, Academe, that focuses on salaries). The National Center for Higher Education Management Systems (NCHEMS) and the federal government's Center for Education Statistics (CES) also have developed classifications of institutions (the classification used in recent editions of CES' Digest of Education Statistics and The Condition of Education is roughly the same as that developed by NCHEMS; an extended version of the latter can be found in Christal, 1986). The advantages of considering this type of comparison group are two-fold: considerable time and effort have already been spent in grouping institutions and the classifications have already established credibility. The problem with these ready-made groups is that they may contain too much within group variation for certain types of comparative analysis. Typically they are based on only a few comparative dimensions such as size and the extent of research activities.

This basic presentation suggests that most institutions already belong to a number of potentially useful comparison groups, and that the issue to be addressed with a comparative analysis determines the appropriate type of comparison group to use. In cases where a comparison group is needed but is not predetermined, a number of techniques exist to help select the proper set of comparison institutions.

DEVELOPING A COMPARISON GROUP

A typology of procedures for developing "peer" groups is described in Brinkman and Teeter (1987) and is displayed in Figure 1. Options for developing comparison groups range from statistical approaches to those that depend entirely on judgment. The top half of the figure describes the continuum of options and the bottom half indicates the techniques themselves (the techniques are meant to be a

representative rather than an exhaustive list). Each of the techniques will be briefly described in terms of the continuum and then examined in more detail.

Cluster analysis and supporting factor-analytic and discriminant techniques are characterized by heavy reliance on multivariate statistics and computer processing. A large number of institutional descriptors can readily be handled. These statistical techniques tend to de-emphasize judgment in the form of administrator (as opposed to statistician-analyst) input. The hybrid approach incorporates a strong emphasis on data and on input from administrators, combined with statistical algorithms for manipulating data. The threshold approach also emphasizes a formal, systematic appeal to data and to administrator input, but it depends little, if at all, on statistical algorithms. Administrator input is heavily emphasized in the panel approach; some data may be included informally but not systematically or comprehensively.

Figure 1

A Typology of Procedures for Developing Peer Groups

- EMPHASIS -			
Data & Statistics	Data & Statistics & Judgment	Data & Judgment	Judgment
Cluster Analysis	Hybrid Approach	Threshold Approach	Panel Review

- TECHNIQUE -

Cluster Analysis and Related Statistical Procedures

Cluster analysis refers to statistical procedures that identify groups of entities that have similar attributes. The purpose of the clustering algorithms is to be sure that entities in a given cluster are more similar regarding the chosen attributes than the entities in any other cluster. The calculation of statistical distance is central to this procedure and alternative ways of making the calculation represent one way of distinguishing alternative clustering algorithms.

Clustering routines can be used with continuous as well as discrete data. The technique does not require the analyst to make arbitrary judgments about the appropriate threshold levels or cut-off points for interval variables included in the analysis. However, there are drawbacks as well. No definitive solutions are provided; judgment

is needed to decide both how and where group boundaries will be drawn and judgment is required in choosing weights to assign to variables entering the analysis.

A variable reduction technique known as factor analysis is sometimes employed as a means of conveniently incorporating a large amount of data in the peer group selection process, often as a prior step before undertaking cluster analysis. Factor analysis is appropriate when several original descriptors are a measure of the same attribute. For example, total enrollment, total degrees awarded, and total instructional expenditures are descriptors whose values are likely to be highly correlated and would generate a factor that would be interpreted as "size." For more complete descriptions about the use of factor analysis to generate groups of similar institutions, see Stephenson (1953), Cole and others (1970), Smart and others (1980), Elsass and Lingenfelter (1980), and Terenzini and others (1980).

Factor analysis also has its limitations, such as the following: sample size directly impacts the reliability of its results, variables are supposed to be normally distributed, judgment is needed to determine the number of factors to extract to create factor scores, and judgment is also required in selecting the rotation procedure. Data standardization in factor analysis can cause problems too. For instance, if the variance of a descriptor such as enrollment in engineering is greater than the variance for enrollment in social science, the engineering variable will have greater impact on the factor results, regardless of whether it is substantively more important for the intended comparison. The same problem can occur when using cluster analysis. In short, there are no right answers mathematically speaking for many of the choices that must be made in using factor analysis.

To examine the results of the clustering technique, one can use discriminant analysis. This technique provides a means of assessing both the "goodness of fit" between institutions and their assigned groups, and the relative statistical influence of variables employed in the cluster analysis. Discriminant analysis is used in this manner in the previously mentioned studies by Elsass and Lingenfelter (1980) and Terenzini and others (1980). Discussion of this procedure can be found in most texts on multivariate statistics.

Cluster analysis is especially useful to the researcher charged with "mapping" the institutional universe (or a large portion of it). Other approaches are probably more suitable when the task is to find a comparison group for a particular institution. (Some computer packages for cluster analysis will output the statistical distance of each institution from every other institution in the sample; one can also use this type of output, a form of proximity analysis, to develop a comparison group for a single institution.)

Hybrid Approach

Various hybrid approaches are conceivable. One such approach is used by the Kansas Board of Regents to identify peer groups for the six four-year institutions under its jurisdiction. There are ten distinct steps in the process. They are briefly described below. A complete description can be found in Cleaver (1981).

1. A subset of states from which peer institutions can be drawn is selected based on criteria such as population, urban/rural mix, and industrialization.

2. Nominal variables are used to create a subset of institutions that ultimately will be rank ordered in terms of similarity to each of the Regents' institutions. To be in that subset, an institution must be public, four-year, and not a branch campus.

3. The subset of all candidate institutions is divided into three groups based on the number of Ph.D. programs offered and the size of the city in which the candidate is located, to correspond with a three-part division of the six Kansas institutions.

4. Campus officials review the institutions in each of the groups and discard any institution whose curriculum is excessively narrow (for example, one with almost all its degrees in engineering or medicine). Any institution that remains in one of the groups is judged to be a viable candidate to be included in the final list of peers.

5. The next task is to arrange the remaining institutions in the order of their similarity with the respective Kansas institution. This involves selecting and assembling detailed information on institutional enrollments, finances, and degrees awarded.

6. The first of a series of statistical algorithms is introduced as the raw data values for the variables assembled in step five are normalized by conversion to z-scores.

7. Comparison scores are created by calculating the difference between each candidate institution's z-scores and its respective Kansas institution.

8. These comparison scores are then standardized.

9. Weights are applied to the comparison scores (the Kansas institutions gave degrees awarded the highest weight, finances the lowest).

10. The standardized and weighted comparison scores are summed to produce a similarity score. The candidate institutions are then rank ordered on these similarity

scores, which concludes the statistical portion of the process.

The rank ordered listings can be used in various ways to support data comparisons.

Threshold Approach

As was true for hybrid approaches, various types of threshold procedures can be envisioned. One such approach, developed at NCHEMS, can be described briefly as follows. The procedure combines raw data, thresholds, weights, and a modest statistical algorithm. Both nominal variables, such as public versus private control, and interval variables, such as enrollment, the number of degree programs, and so on, are used. In the typical application, the nominal variables are used as hard point requirements to reduce the universe of relevant institutions. For example, if "public control" is considered an essential characteristic for a potential peer institution, then any institution not publicly controlled is eliminated from further consideration.

After the nominal variables have been used to generate a subset of institutions, the interval variables are used to rank order the remaining institutions. Points are assigned to each institution based on the importance attached to each interval variable (by means of a simple weighting system) and the number of times an institution misses a prescribed range that is established for each of those variables - these ranges as well as decisions regarding the nominal variables are the thresholds that give this procedure its name. The points are the basis for the rank order: each miss on an important variable pushes the candidate institution down the list further away from the home institution. The rank ordered list is meant to be a guide to analysts and administrators at the home institution who make the final selection of institutions for their comparison group. The last step is an explicit appeal to expert judgment. The same is true for the selection of comparison criteria, their weights, and their acceptable ranges. The procedure is designed to channel and highlight judgment based on data. The transparency of the procedure is a strength but it also can be a vehicle for manipulation. The latter can only be countered by designing appropriate checks and balances into the overall process of selecting a comparison group.

Panel Approach

The "panel approach" uses informed judgment to develop comparison groups. No statistical algorithms are used, and data are used only informally. Consensus among informed and knowledgeable individuals is the basis for the selection of comparison institutions. The appeal of this approach is its simplicity and its adequacy for some comparison purposes such as exchanging policies. Experience suggests, however, that occasionally there may be a need

to replace, or at least validate, the panel approach with objective, statistical approaches. A limitation on its replacement is the lack of widely acceptable quantitative measures of qualitative factors. The ambiguity of quality and other measurement problems that flow from the complexity of various activities in higher education insures its survival.

SUMMARY AND CONCLUSIONS

Institutional comparisons are best begun exactly where any good analysis begins - with a clear sense of purpose. With that in hand, one can address the technical and the human/political dimensions of the comparison process. Both dimensions are important. On the technical side, if an acceptable institutional comparison group already exists, the main area of concern is data. The usual data problems faced by the institutional researcher are almost always compounded in a comparative analysis. The best approach is to proceed with caution, assuming as little as possible about the quality of the data. If the task includes developing a comparison group, there are a number of procedures available. Choose one that suits your analytic skills, the purpose of the comparison, and the broader political aspects of the task. To have utility, the procedure must appear reasonable and valid to those who are to use its results. Involving interested parties in the process often helps, especially if the involvement is well managed and occurs early when purpose is specified and selection criteria are chosen.

REFERENCES

Brinkman, P. T. & Krakower, J. (1983). Comparative data for administrators in higher education. Boulder, CO: National Center for Higher Education Management Systems.

Brinkman, P. T. & Teeter, D. J. (1987). Methods for selecting comparison groups. In P. T. Brinkman (Issue Ed.), Comparative and interinstitutional analysis. New Directions for Institutional Research (No. 53). San Francisco: Jossey-Bass.

Carnegie Commission on Higher Education (1976). A classification of institutions of higher education. Berkeley, CA: Carnegie Commission on Higher Education.

Christal, M. E. (1986). Higher education financing in the fifty states: Interstate comparisons fiscal year 1984. 5th ed. Boulder, CO: National Center for Higher Education Management Systems.

Christal, M. E. & Wittstruck, J. R. (1987). Sources of comparative data. In P. T. Brinkman (Issue Ed.), Comparative and interinstitutional analysis. New Di-

rections for Institutional Research (No. 53). San Francisco: Jossey-Bass.

Cleaver, G. S. (1981). Analysis to determine a ranking in similarity for institutions in higher education. Paper presented at the Annual Conference of the Society for College and University Planning, Omaha.

Cole, N. A., & Others (1970). An analysis of spatial configuration and its application to research in higher education. ACT Research Report, No. 35. Iowa City, IA: American College Testing Program.

Elsass, J. E. & Lingenfelter, P. E. (1980). An identification of college and university peer groups. Springfield, IL: Illinois Board of Education.

Jones, D. P. (1982). Data and information for executive decisions in higher education. Boulder, CO: National Center for Higher Education Management Systems.

Smart, J. E., & Others (1980). Qualitative and conventional indices of benchmark institutions. Paper presented at the 20th Annual Forum of the Association for Institutional Research, Atlanta.

Stephenson, W. (1953). The study of behavior. Chicago: University of Chicago Press.

Teeter, D. J. (1983). The politics of comparing data with other institutions. In J. W. Firnberg & W. F. Lasher (Issue Eds.), The politics and pragmatics of institutional research. New Directions for Institutional Research (No. 38). San Francisco: Jossey-Bass.

Terenzini, P. T., & others (1980). A conceptual and methodological approach to the identification of peer institutions. Research in Higher Education, 12, 347-364.

THE INSTITUTIONAL FACTBOOK: KEY TO PERCEPTION OF INSTITUTIONAL RESEARCH AND INFORMATION DISSEMINATION ON THE CAMPUS

James O. Nichols
University of Mississippi

Richard D. Howard
North Carolina State University

Bobby H. Sharp
University of Mississippi

The Chief Executive Officer at your institutio has just asked for an overview of the institutional re search office's activities and contributions to the institution, as well as samples of your work in connection with an ongoing assessment of institutional operations and administrative costs.....The Executive Secretary of the Board of Trustees has just called and asked that you forward to each of the three finalists for the vacant Chief Executive Officer's position a summary of the quantitative aspects of your institution's operations. Your response to his request will constitute the first impression which your new boss will have of your work.....Your institution has just launched a strategic planning process and in the midst of an initial meeting of the institutional planning committee, your Chief Executive Officer asks "What do we have that can serve as a common point of reference or description of current operations from which we can begin considering the future?".....Your institution has just begun its institutional self-study for reaffirmation of accreditation and the Director of the Self-Study drops by your office and asks "What do we have in the way of a statistical description of the institution that can be used by our campus committees and the visitation committee from the accrediting Agency?.....In each of these cases, what are you going to provide? The answer may be your institutional factbook.

What then is an institutional factbook which might answer the variety of questions posed above? An institutional factbook can be described as an official and recurring publication (in one form or another) which consists primarily (though not exclusively) of a compilation of statistical information designed to describe the quantitative aspects of institutional operations.

FACTBOOK PURPOSES

The purposes which an institutional factbook can serve are best considered from the perspectives of (a) what it can do for the campus and (b) what it can do for the institutional research component. There are at least three purposes which an institutional factbook can fulfill on the campus:

1. Historical Record: In the short-run, a factbook can serve as the official source of data for the completion in a consistent manner of questionnaires by various agencies on campus. Over the longer perspective, this role becomes somewhat archival in nature as the compilation of data in a consistent format each year traces the development of the institution. The establishment of this official historical record also answers the question of what information may be released to the public.

2. Public Relations: Some factbooks are intentionally produced with the primary purpose of influencing the reader, more than informing him/her. Such publications select and highlight certain information intended to support a particular institutional case or position. The intended audience for these factbooks is frequently (though not always) off-campus constituents such as those who provide funds and board members.

3. Decision/Planning Support: Factbooks intended to support either current or future decisions on the part of the institution differ in both content and candor from those published for historical or public relations purposes. Decision/Planning Support type factbooks tend to focus on different data (teaching loads, average salaries, etc.) than primarily historical publications and frequently contain more sensitive information than factbooks prepared primarily for public relations find comfortable.

From the standpoint of the institutional research component, a factbook can serve several purposes:

1. Institution-wide demonstration that institutional research is taking place. On many campuses, the impact of the institutional research component is primarily limited to senior administrative/academic personnel. Publication and dissemination of a factbook demonstrates to much of the rest of the campus that the institutional research component is indeed active.

2. Insurance that important data reach a wide audience. It is unfortunate, but important, to note that some of our administrative/academic colleagues on each campus may restrict dissemination of important information forwarded to them in order not to compromise policy positions which they have taken. The relatively widespread dissemination of a factbook directly to the departmental level insures that those most likely to be impacted by the institutional research product also benefit from its availability.

3. Lead the reader into more detailed institutional research studies. While the institutional factbook will not be able to contain detailed, discipline-specific information resulting from some institutional research activities, institution-wide summaries of such studies

published as examples in the document can serve to lead readers to the results of these studies which directly impact themselves.

What then are the disadvantages to publication of an institutional factbook? The primary disadvantages relate to time and money. Because there is no pressing "requirement" that an institutional factbook be published, a conscious value judgement must be made to commit limited resources in the institutional research component to this purpose, whereas other "required" functions such as external reporting may automatically assume a higher priority. Depending on the level of sophistication desired and availability of data, preparation of factbook copy may take between a single person week and several person months. Again, dependent on the size, sophistication, and number of copies desired, the out-of-pocket cost (exclusive of time on the part of office staff) for factbook publication may range from several hundred to several thousand dollars annually.

Considered objectively, a substantial amount of cost/effort goes into production of information by the institutional research component and in other offices at the institution. This information frequently fails to achieve its full usefulness on the campus because "the right people" are not aware of its existence. In order to fulfill any of the institutional purposes identified previously, at only a modest additional cost compared to that already expended to produce the data, numerous institutions will continue to make the decision to publish a factbook.

DECISIONS IN FACTBOOK DESIGN

The decision having been reached to publish an institutional factbook, many of the strategic decisions regarding the publication flow logically from the document's intended purpose (see Figure 1). While no factbook serves any one of the purposes described exclusively, an early decision concerning the primary purpose which an institution's factbook should serve will focus its development and materially assist in making the strategic decisions concerning publication of the document outlined in Figure 1.

In addition to those "Strategic Factbook Decisions" outlined on Figure 1, early determination of the form (bound booklet, looseleaf, or potentially electronic), distribution (number of copies), data presentation modes (tabular, graphic, and/or narrative), and page orientation (upright or horizontal) is desirable. Each of these decisions impacts the selection and compilation of data, as well as the ultimate cost of the document.

Figure 1

IMPACT OF FACTBOOK'S INSTITUTIONAL PURPOSE
ON STRATEGIC DECISIONS CONCERNING ITS PUBLICATION

STRATEGIC FACTBOOK DECISIONS	TYPES OF FACTBOOK INSTITUTIONAL PURPOSES		
	HISTORICAL RECORD	PUBLIC RELATIONS	DECISION/PLANNING SUPPORT
AUDIENCE	.Primarily internal administrative/ technical personnel	.External Constituents- Legislators, Board Members .Internal Constituents- Faculty, other power blocks	.Internal Decision Makers- Chief Executive, Vice Presidents, Deans, Department Chairpersons
TIMING OF PUBLICATION	.Close of fiscal/academic year .Date of publication not critical	.Frequently early fall to include early semester data to influence constituents .Timing critical-rush to publish to influence decisions	.Either close of fiscal year or fall semester .Timing not critical assuming earlier piecemeal distribution of contents
TYPE OF DATA CONTAINED	."Counts" of routine nature-students, faculty, degrees, etc.	.Relatively simple and often preliminary data easily understood by lay persons .Strictly "good news" type data	.Focus upon data required to support decisions - teaching loads, salary patterns, etc. .Contains mixture of good and bad news type data
LEVEL OF DATA	.Institutional and major units (Colleges/Schools)	.Primarily institutional level for external consumption .Occasional aggregation at School/College to make specific point	.From institutional to departmental to support decision making .References to more detailed data available separately
MEANS OF PRESENTATION	.Primarily tabular	.Extensive use of graphics to convey message	.Mixture of tabular, graphic, and narrative presentation

DATA CONTAINED IN INSTITUTIONAL FACTBOOKS

The selection of data for the publication will be heavily dependent on the primary purpose which is identified for the document, but also should relate to the important issues facing the institution at the time. As examples, one would expect the institutional factbook at a private institution with a declining enrollment to focus more upon factors influencing student selection of the institution and relative student tuition or cost of attendance, than would such a publication at a rapidly growing public institution whose factbook might focus on appropriated support per FTE student and classroom utilization.

In general, institutional factbooks convey information concerning students, academic operations, finances, and facilities. Information contained about students frequently can be categorized as concerning student admissions, current enrollment, and future enrollment prospects. Data concerning entering student test scores, geographic origin, planned educational majors, and family income are frequently provided. Information presented about currently enrolled students often includes summaries of enrollment by student level, sex, school/college, and, in an increasing number of factbooks, degree programs. Those factbooks supporting planning also frequently include information concerning future enrollment prospects (projections) based upon either specific or various sets of assumptions.

The academic operations of the institution are usually described with data concerning student enrollment in classes, faculty characteristics, student attrition/retention, degree conferral, and combinations of these areas. Information concerning numbers of student credit hours by department and school/college is the most common representation of class enrollment. In virtually all factbooks, one can find tabulations of faculty by academic rank and highest degree held. In those factbooks designed to support decision making/planning, it is also not uncommon to find data concerning tenure rates and average faculty salaries by department. In addition to these single source type data, sections of the publication concerning academic operations may include information resulting from combination of data from several sources such as teaching loads, instructional cost per credit hour, and degree program cost.

Factbook sections concerning finances and facilities are frequently more modest in scope than the sections outlined earlier. Financial information provided normally consists of expenditure and revenue summaries covering recent years. Physical facility data are often limited to classroom utilization, dormitory occupancy, and building size/capacity.

Less frequently, but occasionally, found in institutional factbooks are official tables outlining the

administrative/academic organization of the institution, listings of administrative officers and members of the Board of Trustees (or its equivalent), and a brief history of the institution.

Following general identification of the data to be contained in the publication, it will be appropriate on many campuses for the institutional researcher to review the data to be contained with a more senior administrative officer to obtain a "blessing" for relatively widespread release of the data. What appears to be innocuous to the novice institutional researcher may, in the context of his/her institution, be very sensitive.

DATA PRESENTATION AND FACTBOOK PUBLICATION

These greater issues having been determined, the institutional researcher must turn to other matters related to data presentation and production of the publication. Among those decisions is the question of data depth or "How many years into the past should the data presented be extended?" One immediate response to this question is that the more depth of data the better. On the other hand, only so many columns of data can be contained in a table, and a realistic decision must be made. This decision will naturally vary from page to page and subject to subject; however, one general rule which can be useful in making this decision is not to include more depth than will be useful in analysis of the data presented. Hence, if an institution has undergone an enrollment decline for the past six years, information concerning enrollment for at least the past six years should be included in the factbook. Only to the extent convenient, however, should data from years prior to that time be included.

The means for presentation of data in institutional factbooks can be categorized as tabular, narrative, and graphic. Tabular data presentations include the provision of data in tables of all types, while narrative presentations include prose summaries or highlights which relate data in sentences or paragraphs. Finally, graphic presentation of data includes the use of all types of charts, graphs, or other pictorial techniques. By far the most common of these forms found in institutional factbooks is the tabular presentation of data.

Among the advantages of tabular presentation of data are:

- familiarity of institutional researchers with the form,
- ease of preparation, and
- efficiency in conveying maximum information in minimum space.

Unfortunately, tabular data presentation exhibits just as many disadvantages among which are its:

- relatively uninteresting or unappealing esthetic qualities,
- difficulty in drawing attention to important highlights, and
- frequently repetitious and "boring" use.

While the tabular form of data presentation will undoubtedly remain the primary means utilized for data presentation in institutional factbooks, the following suggestions are made concerning its use.

1. Vary the format (spacing, margins, etc.) utilized so as to reduce its repetitious nature and provide each table a "personality" of its own.
2. If in doubt, include less, rather than more, detail in tabular data.
3. Intersperse graphic and narrative presentations to "break up" the repetitive nature of multiple, consecutive tables.
4. Use graphics and narrative to highlight or summarize the most important aspects shown in tabular presentations.

Long ago it was stated that one picture is worth 10,000 words. For highlighting a factbook's most important findings, one graphic is worth 20 tables. Graphic presentations can capture the readers' imagination, illustrate relationships between data, and bring life to publications which otherwise may go unread. Tufte's work referred to at the close of this chapter is an excellent non-technical source regarding this subject. Unfortunately, these same graphical presentations are frequently expensive, time consuming to produce, and inherently inefficient in conveying the large amount of data contained in most factbooks.

Because of its strengths and limitations outlined above, the following suggestions are offered concerning the use of graphics in institutional factbooks.

1. Carefully plan the use of graphics to highlight only the most important data.
2. Restrict the amount of data conveyed through any graphic to several key points.
3. Provide tabular data in which the reader may gain further understanding of the information conveyed by the graphic.
4. Avoid presentation of one graphic after another.

Often overlooked in factbook data presentation is the significant value which can be gained through the use of the narrative presentation of data. Along with graphical data presentation, the preparation of narrative highlights drawn from subsequently presented data in tabular format is among the most effective techniques designed to lead toward greater reader interest in the data presented. Narrative passages are also extremely useful to explain complex tabular presentations and examples which might otherwise be misunderstood. Institutional factbooks will always remain primarily tabular in nature; however, documents that utilize graphic and narrative data presentations to highlight and vary this overall presentation probably are the most widely read and hence successful publications.

USING COMPUTERS IN FACTBOOK PREPARATION

Factbook preparation offers numerous opportunities for exploiting the power of computers and computing technology. As tools especially suited to organizing, storing, and retrieving data, computers can facilitate all phases of factbook preparation, from the initial "data crunching" to the finished form of data presentation.

Processing large amounts of raw data with computers is common to most institutional research offices. Data are selected, sorted, and summarized, and calculations are performed by computers as typical institutional research functions. Most of these applications traditionally have been handled on mini or mainframe computer systems with structured programming languages, but as micro computers and packaged software have become more powerful, an increasing number of statistical packages, electronic spreadsheets, and data base systems are frequently used to process the mass of data gathered and maintained at most institutions.

Since a primary factbook function is to condense into a practical volume a large amount of data related to an institution, the task can be greatly facilitated by using computers to design, store, retrieve, and update text, tables, and graphics. Among the specific uses are the following.

1. Word Processing. With space a critical commodity in a factbook, text placement and appearance become very important considerations in factbook preparation. Word processing not only facilitates the initial entry and design of text in a factbook, it also allows for easy revisions and reorganization as the factbook is edited for publication.

   A second benefit of using word processing in factbook preparation occurs when subsequent issues of a factbook are produced. Pages from previous issues can be retrieved, easily revised, and printed as camera-ready copy for publication.

2. Electronic Spreadsheets. Factbooks typically contain many tables. An electronic spreadsheet greatly simplifies the task of designing and revising tables in order to maximize space use and to present data in an esthetic manner. This use of an electronic spreadsheet typically does not exploit the analytic power of such software, but often can be used to calculate and present summary statistics and percentages.

   Electronic spreadsheets also allow for easy retrieval and revision of previous factbooks. For example, updating a table in an annual factbook may consist of simply deleting the first year's data. Sums, averages, and percentages remain in place and are automatically calculated reflecting the new data. Revised tables are then ready for printing camera-ready copies.

3. Graphics. Computer generated graphics have become more sophisticated in recent years, enhancing the ease with which information can be depicted pictorially in factbooks. Mims' recent monograph concerning computer graphics is a good source regarding this subject. Factbook graphics that were once produced by graphic artists can be readily created, stored, and revised as needed. Data showing trends or contrasts among items or groups lend themselves well to graphic portrayals.

4. Electronic Factbooks. Some institutions have created and maintained facsimiles of their factbooks "on-line" within their institutional computer systems, accessible through computer terminals. Printing costs are thus reduced and revisions are simplified. More institutions may adopt on-line factbooks as campus computer systems and networks become more advanced.

Computers have long been essential tools used by institutional researchers and in factbook preparation. The applications have changed, however, as micro computing technology has introduced powerful, easy-to-use word processing, electronic spreadsheet, data-base management, and graphics software. Both increased productivity and better designed factbooks can result from applying current computer technology to factbook preparation.

FACTBOOK LIFECYCLE

While much attention will be focused upon the actions necessary to publish the initial edition of an institutional factbook, it is important to remember that the factbook is a recurring publication. In that regard, adaptation and refinement of the original edition become primary considerations in subsequent editions which fortunately require less effort.

There probably has never been an original edition of an institutional factbook series with which its authors were completely satisfied. The desire to include one more ta-

ble, another graphic, or a further word of explanation is typical of the feelings expressed by the authors of initial factbook editions. Future editions provide the opportunity to make these adjustments and those responsible would do well to make notes at the time of original edition release of those things which they would like to have done differently and make notes as well during the year as suggested changes are incorporated.

Because a factbook should focus upon major campus issues which change (at least in degree) from year to year, it will be necessary to review the content annually and make necessary adjustments to adapt the focus of the publication to the then current campus environment. Adjustments of this type are relatively minor when viewed on an annual basis (add a table or graphic or two, delete no longer needed data, etc.), but when viewed over several years can materially impact the composition of a factbook series.

One welcomed aspect of subsequent editions of a factbook is the much reduced burden which its publication brings about. Assuming use of micro computer capability, tabular presentations occupy approximately 10-20 percent as much time in subsequent presentations as in the original edition. Graphic efforts are also greatly reduced as adjustments to original graphic presentations replace complete design. Whether micro computer generated or done manually, always retain the original copy from which any edition is printed, as it represents a good start on next year's edition.

Factbooks, like most projects, exhibit life cycle characteristics. Just as they are born amid great fanfare, they should ultimately go through a period of evaluation and subsequent renewal or discontinuation. Every five to ten years, even the best factbooks should be examined to justify their continuation. In this regard, the institutional research component should query its constituents periodically concerning the usefulness of the publication, suggestions for its change, or possible discontinuation. Armed with these responses, the institutional research component should carefully review, refine, or discontinue the publication.

FACTBOOK EXAMPLES

Where can one look for examples of institutional factbooks? There are at least three likely sources for examples of institutional factbooks. Each year at the annual Forum of the Association for Institutional Research the "Publications Display" contains examples of the best institutional factbooks published throughout the country. Second, the Educational Resources Information Center (ERIC) Clearinghouse on Higher Education has included in its collection numerous examples of institutional factbooks and these documents are widely available on microfilm in libraries on many campuses. Finally, the factbooks published by several institutions, including the University of Mississippi and Georgia State University, have been nationally recognized and copies of their publications may be requested from those institutions.

FUTURE OF INSTITUTIONAL FACTBOOKS

While technology will undoubtedly continue to change the means through which factbooks are prepared, the basic purpose which such publications serve will not be altered. Factbook type publications will remain among the most effective and efficient means for the widespread dissemination of a recurring basic compilation of information concerning institutional operations. While some institutions may convert to an on-line computer accessed set of data serving as their institutional factbook, use of micro computers to ease the burden of publication probably assures that the vast majority of institutional factbooks will continue in hard copy format. While the institutional factbook is among the least sophisticated and oldest of the tools available to the institutional researcher, its key role in development of the institutional research component and dissemination of its product appears to be assured well into the future.

REFERENCES

Bowman, W. J. (1968). Graphic Communication. New York: Wiley.

Dickerson, G. C. (1973). Statistical Mapping and the Presentation of Statistics. New York: Crane, Russak and Company.

Lasher, W. F. (1978). The comprehensive institutional planning process and the role of information in it. Planning for Higher Education, 6 (4),11-15.

Leischuck, G. S. (1970). Communicating the results of institutional research: The production and users of the factbook. In P. S. Wright (Ed.), Institutional Research and Communication in Higher Education Tallahassee, FL: Association for Institutional Research.

Mims, R. S. (in press). Handbook on Micro Computer Graphics (Tentative Title). Tallahassee: Association for Institutional Research.

Smith, G. (1980a). Systematic information sharing in participative university management. Journal of Higher Education, 51 (5), 519-526.

Smith, G. (1983). Institutional factbook: Catalyst for an integrated institutional research program. In N.P. Uhl (Ed.) Using Research for Strategic Planning, New Directions for Institutional Research No. 37. San Francisco: Jossey-Bass.

Tufte, E. R. (1985). The Visual Display of Quantitative Information. Cheshire, CT: Graphics Press.

USING STATISTICAL PACKAGES/SPREADSHEETS

Bernard D. Yancey
University of Texas-Austin

Maryann S. Ruddock
University of Texas System

As the raw materials of a carpenter are wood and other building materials and the process is construction, the raw materials of the institutional researcher are data and the process, data analysis. Each trade has its own set of tools. While the tools of a carpenter are his hammers, saws, chisels, etc., the tools of the institutional researcher are his calculators, computers, and software packages. Just as a good carpenter must learn which tool to use at which stage of the construction process, the institutional researcher must learn which tools are best applied in which situations and at which stages of the data analysis process.

For the institutional researcher, the various stages of the data analysis process can be simply stated as follows.

1. Identification. Identifying the data and data sources.

2 Acquisition. Acquiring the needed data.

3. Conversion. Converting the data to a usable, usually computer readable format.

4. Auditing. Auditing the data to determine its accuracy, validity, and reliability.

5. Preliminary Analysis. Determining what the characteristics of the data are so that additional analyses can be chosen.

6. Analysis. Analyzing the data in light of the questions that have been raised and the results of the preliminary analyses.

7. Summary. Summarizing the results of the analyses.

8. Presentation. Presenting and interpreting the analyses' results.

Statistical packages and spreadsheets are simply tools that are available to the institutional researcher. While these tools can primarily be used at steps 4 through 8 of the data analysis process, some statistical packages, such as SAS[1] and SPSS-PC[2] contain components that can greatly aid in step 3 of the process, data conversion. Spreadsheets can also be useful at this stage if the initial data conversion is to occur using a microcomputer.

When considering statistical packages and spreadsheets, deciding which to use and when is not always easy. There are a variety of factors that must be considered, including the type of calculations to be performed, the quantity and quality of the data and the skill level of the user. It may also be that a particular problem can best be approached by using both statistical packages and spreadsheets. One important point, however, should always be kept in mind. The use of such packages provides no panacea. Computers and such software packages are no substitutes for organization and a thorough understanding of the procedures and processes involved. While computers can produce results perhaps a thousand times faster than by hand, they produce and compound errors just as rapidly.

Before continuing, it is perhaps best to describe briefly what is meant by the terms "statistical package" and "spreadsheet". A comprehensive description of both terms is simply not possible here, but the following will perhaps aid the reader. The initial purpose of the spreadsheet was to provide a means of manipulating and making calculations on data that are stored and arranged in labeled columns as would be found in a ledger. The initial spreadsheets were designed to aid the business community with repetitive and often labor-intensive calculations. Spreadsheets were also designed to allow management to make repeated calculations and to explore "what if" situations without having to rely on large mainframes. The rectangular data structure of spreadsheets makes it quite easy to perform repeated calculations across rows or columns, particularly if these calculations are to be made for the whole data set. If multiple passes must be made on the data with intermediate results used in subsequent analyses, difficulties arise. The capabilities of current spreadsheets have been greatly expanded, but they still embody the general philosophy that led to their initial development.

Statistical packages, on the other hand, were initially and specifically designed to perform statistical analysis. Such packages were developed to calculate a variety of statistics, both simple and complex, that were needed to support hypothesis testing. While the rectangular data structure used by most statistical packages is similar to that found in spreadsheets, the level and magnitude of the computations are considerably more complex. While spreadsheets generally are limited to making calculations based at most on columns or rows, statistical packages generally perform calculations on matrices. A number of the statistical analyses require making multiple passes on the data with the results from earlier passes being used in the subsequent stages of the analysis.

The material to be presented here will first briefly discuss some of the factors that need to be considered in determining whether or not a statistical package and/or a spreadsheet would be appropriate. A brief discussion will then be provided of the various points to be considered

in selecting a statistical package, along with a discussion of some of the strengths and weaknesses of some of the major statistical packages on both mainframes and microcomputers. A similar approach will be taken for spreadsheets. In addition, a brief discussion will be provided describing how to integrate statistical packages and spreadsheets.

SPREADSHEETS, STATISTICAL PACKAGES, OR BOTH?

There are no hard and fast rules that can or should be followed for determining whether to use a spreadsheet, statistical package, or both. It is not simply a matter of comparing the technical capabilities of a given statistical package or spreadsheet. The intangible factors that must always be considered are the skills, experience, and technical abilities of the researcher who will be using the spreadsheet or statistical package. However, the following points may prove helpful.

Efficiency: One important factor that should always be considered is efficiency, not only in terms of machine resources and computing costs, but also the time and commitment, and in particular the salary costs of the individuals conducting the analyses. In general, the question to be answered should be what tool or combination of tools will let you do a quality job with the least total expenditure of resources? For example, one package may cost more in terms of computer time to perform the same analysis, but it may take only one tenth the man hours to set up and run. With the decreasing cost of computing resources, the major cost in performing analysis is often salaries. If the overall picture is considered, a $10.00 saving in computer costs may produce a $90.00 increase in hourly costs.

Availability: An important factor may simply be availability. The researcher may already have access to a spreadsheet program on a microcomputer, but not to a statistical package. Most institutional researchers, however, will have access to one or more statistical packages. Access to statistical packages on mainframes can, however, exist in varying degrees. A particular statistical package may reside on a mainframe, but the machine may be so heavily used or restricted that the researcher can only get access for limited time periods, or it may take from 24 to 36 hours to turnaround a job. It may also be that while a particular package is available, there may be limited local support in terms of documentation or consulting. The cost of mainframe computing may also simply be prohibitively expensive, or it may be that the data is of a confidential nature and the researcher is limited to using software that exists on a secure machine.

User Expertise: This is perhaps one of the most intangible and yet important points to be considered. The expertise and ability of the user should be considered along two lines. The first and least important is the user's skill

level for interfacing with the computer hardware and software. A more important consideration is the user's knowledge and experience in the area of statistics and statistical analysis. Access to a statistical package and a "cookbook" does not instantly make a researcher a statistician. In recent years an increasing number of journal articles have been published detailing rather sophisticated statistical analysis. These analyses involve complicated and time consuming calculations. A number of the commonly used statistical packages have only recently added routines that make these analyses easier to do. One can only wonder what proportion of the analyses that are being seen are being performed by knowledgeable researchers who have been waiting for these labor saving tools and what proportion are being performed by researchers who do not have either the sufficient background or experience to understand the underlying statistical process and assumptions. More to the point, are these analyses now being performed simply because they are easy to do?

MATCHING THE TOOLS TO THE JOB

If the researcher has access to both spreadsheets and statistical packages, then he must decide how to use these resources. In this situation, the following should be considered.

The Nature of the Analyses: Are the analyses to be performed descriptive or inferential? If the analyses are strictly descriptive, depending on the quality and quantity of the data, it may be that a spreadsheet could be used. If, however, any form of hypothesis testing is being considered and probability values are required, it is quite likely that a statistical package should be considered.

The Unit of Analysis: If the unit of analysis is the individual and simple repetitive calculations are to be performed for each individual with the results being displayed for each individual, and if there are not too many cases, it is quite likely that a spreadsheet will fill the bill.

Data Quantity: If only simple descriptive statistics (means, standard deviations, counts, etc.) are needed and the data will fit on one page, then a spreadsheet may also prove useful. One of the inherent weaknesses of spreadsheets is the inability to easily produce frequency distributions and/or cross tabulation tables, particularly for large data sets. In addition, as the number of variables and the number of cases increase, analysis using a spreadsheet becomes more and more difficult and time consuming. With a statistical package, if the number of cases and/or variables increases, it may simply be a matter of waiting a little longer for the analysis run to finish.

What If Questions: Spreadsheets are particularly useful for planning and "what if" analyses such as budgeting and resource allocation. For example, consider a situation

where financial aid of varying amounts is awarded to Master's, Ph.D.'s, and continuing students. If there is a fixed amount of total aid, then it can be useful to determine how various mixtures of aid packages will fit in the total fixed amount of aid.

Multiple Group Comparisons: In situations where two or more groups of individuals are to be compared and more than simple percentages are required, it is often more efficient to use a statistical package. This is particularly the case when probability values are needed. While it may be technically feasible to do an Analysis of Variance using a spreadsheet, the effort involved could be considerable, depending on the complexity of the analysis.

Maybe you need both: The data analysis process is made up of multiple steps. The needs of the researcher at each step can vary considerably, depending not only on the questions that have been raised, but also on the quantity, quality, and characteristics of the data and the skill and experience of the researcher. It may be that the optimal solution involves the combined use of statistical packages and spreadsheets. Most commonly used statistical packages and spreadsheets allow for the interchange of data in a variety of formats. It is no longer an insurmountable task to transfer data from one spreadsheet to another or from a spreadsheet to a statistical package.

STATISTICAL PACKAGES

Statistical packages and spreadsheets are tools. The goal is to match the tools with the job. Following this line of reasoning, when the researcher has more than one statistical package available, the problem becomes one of determining how these packages are to be used. One common mistake made by researchers is to learn only one package and then try to perform all the analyses with that one package. The more tools a researcher has available, the more likely that an efficient and useful tool can be found. Statistical packages on both mainframes and micros are becoming easier and easier to use. The old excuse that a researcher simply does not have time to learn more than one package no longer holds. Most of the available packages are based on similar concepts and it is becoming easier to switch from one package to another and to transfer data between packages.

MAJOR POINTS TO CONSIDER IN CHOOSING AMONG STATISTICAL PACKAGES

Auditing and Preliminary Analyses: Contrary to popular belief, one of the first things that needs to be considered in choosing a statistical package is not simply the particular analyses that are to be run, but rather how much data manipulation, data auditing, and preliminary analysis must be performed. Most of the commonly used statistical packages provide a wide range of statistical procedures that will meet the needs of most researchers.

Most practitioners in the field of institutional research realize that about 90 percent of the time and effort spent in doing data analysis involves getting the data ready to run and only about 10 percent in actually running the analyses. What is primarily needed is not just a statistical package, but a data management system. These data management capabilities do not necessarily have to reside in the statistical package itself, but do need to be available. It may be that they can be performed by using a database management system in conjunction with a statistical package. It is, however, extremely convenient to have the data management capabilities in the statistical package itself.

The most often overlooked part of the analysis process is data auditing and preliminary analysis. With respect to data auditing, the quality of the analyses and the validity of the results are directly related to the quality and validity of the data. The old saying of "Garbage in, Garbage out" is particularly relevant in this situation. This data auditing process should at minimum consist of checking for out-of-range values and in most instances should include checks of the internal consistency and integrity of the data. Such audits require first and foremost the liberal and judicious application of common sense. Take nothing for granted. For example, it would be at least embarrassing to report headcount data listing three males residing in what is supposed to be a women's dorm, or mean test scores that are higher than the maximum possible score for a test.

Before statistical analysis procedures can be chosen, the characteristics of the data to be analyzed must be determined. This involves assessing not only the quality and quantity of the data, but also its statistical characteristics. Are the values of the variables that are to be analyzed normally distributed? How widely dispersed are the values of the various variables? Are the frequency distributions symmetric? Almost all statistical analysis procedures are based on a specific set of assumptions. For example, a considerable number of the classical methods assume normality or, at minimum, homogeneity of variance. It should be remembered that the calculation of a statistic within a given procedure is a mechanical process that can be accomplished in most instances regardless of whether the underlying assumptions are satisfied or not. Matching the statistical procedure with the underlying assumptions is sometimes referred to as finding an appropriate statistic. Technically speaking, however, there is really no such thing as an appropriate or inappropriate statistic. What is appropriate or inappropriate is the interpretation of that statistic. If the underlying assumptions are not met, then, while a statistic can be calculated, it may be meaningless, or, unfortunately, in too many instances totally misleading.

Before the researcher can choose a statistical procedure that will produce meaningful and useful results, the re-

searcher must not only consider the specific questions that have been raised, but fully understand the quality, quantity, and statistical characteristics of the data that is to be analyzed. It is in this process that the quality statistical package can most readily prove its worth.

One important feature of any software package that is used at this step is whether or not it has interactive capabilities. Preliminary analysis may require as many as 100 or more separate analyses. If the package being used is interactive and can provide immediate feedback, this process can be greatly shortened. This will also let the researcher perform additional analyses that may not have been considered using a non-interactive system.

Once the basic characteristics of the data are understood, the researcher is ready to consider the following in choosing a particular package.

One Analysis or a Series: Is only one analysis to be performed or will a series of analyses be needed? If only one analysis is to be performed, then the key question is whether or not a package can perform the analysis that is needed. The task at hand may, however, require a series of analyses with the results of the earlier ones used in subsequent analyses. In this situation, a statistical package that allows for saving intermediate results, usually in the form of output files, can be extremely useful. Given the data interchange capabilities of more statistical packages, it should not always be considered a requirement that all of the analyses to be performed can be performed by one package.

Skill Levels: What are the skill levels of the individuals performing the analyses? Depending on the individual, some packages are simply easier to use than others. It may also be that the individual performing the analysis has had more experience with one particular package. In this situation, the start-up time for learning a new package should be considered. How comfortable a researcher feels with a given package is also something that should be considered. How knowledgeable is the researcher who will be performing the analysis in the area of statistical analysis, design and interpretation? Will he be able to determine whether or not the data meets the underlying assumptions for the analyses that are to be performed? If the data and/or the questions to be answered do not exactly fit one of the standard designs, can he determine which input parameters need to be changed and make the appropriate changes?

Available Support: How much local support is available? Are consultants available locally to help with questions? Is documentation available? Are there researchers on campus or in your department who have successfully used a particular package? What do you do if no one locally can answer your questions? Is there a national hotline or information source that can be tapped? Where do you go for training?

Time Constraints: How much time is available to conduct the analyses? In the area of institutional research, analyses results are often useful only if they are timely. Some packages and approaches simply take longer to produce results than others.

The Quality of the Output: How readable and usable is the output? If the output is in a form that is not easily read or interpreted, then the extra time that will be needed to interpret and use the results should be a factor considered in the deliberations. Some statistical packages will produce analysis results that can be directly used with only minor modifications in the final report. Some packages also produce more output than can ever really be used. It is possible that the inexperienced researcher may simply pick and choose from the multitude of statistics that are provided in such situations and never consider that each statistic requires that a specific set of assumptions be satisfied before they can be meaningfully interpreted.

Efficiency and Costs: What will the computing costs be, both in terms of real dollars and computer resources? Some statistical packages cost more to run and use considerably more resources. When estimating cost there is a tendency among computer operators and systems programmers to look only at computer resource requirements. Usually they totally ignore the salary costs of the individuals using a package. The whole picture must be considered.

USING STATISTICAL PACKAGES

An institutional researcher who acquires a copy of the manual for a statistical package and proceeds to perform the analyses using a "cookbook" approach without a basic understanding of the process and procedures involved, and particularly underlying assumptions and limitations, is doing great disservice not only to himself, but also to the profession. A review of the recent literature reveals an increasing use of rather sophisticated multivariate analysis techniques. While the techniques can be extremely useful, they are sometimes applied to problems that could have as easily been addressed using less sophisticated and easier to interpret techniques. It is always a good idea to try to use the simplest analysis technique that will adequately address the questions that have been raised.

General Guidelines for Using Statistical Packages

Take nothing for granted. The first and foremost point to be considered when using statistical packages, or for that fact, any computer software, is to take NOTHING for granted. If the results for a particular analysis run look strange, in all likelihood they are the result of some misstep in the analysis process; the wrong parameter on a control card, incorrectly coded or missing data, or the wrong model, etc.

When you first start using a particular statistical procedure, it is always a good idea to benchmark the procedure. This can be done by analyzing a data set from a statistical text and comparing the results with the results given in the text book. This is particularly important with some of the more sophisticated statistical procedures. For example, from the descriptions given in the manuals, it is often difficult to determine just what model is being used in some of the Analysis of Variance or Analysis of Covariance procedures. The interpretation of an Analysis of Covariance can vary widely depending on whether the model assumes parallel or separate slopes.

Another way to benchmark a new statistical package procedure is to perform the same analysis using two different packages and compare the results. Different packages may use different algorithms for obtaining the same results. These algorithms may have a subtle and yet meaningful impact on the results.

Consider the whole process. Statistical analysis does not end when the analyses are performed. The results must be interpreted and summarized in a format that can easily be presented and understood. Statistical packages that produce results in a format that simplifies this process can be extremely useful. Several of the more widely used statistical packages (SAS, SPSS-X[2], and BMDP[3]) can produce output which can often be inserted into final reports with only minor modifications. The PROC TABULATE feature of SAS and the report generator and TABLES routine in SPSS-X can be particularly useful. One way of effectively using the output from these procedures would be as follows. The analysis can be performed on a mainframe and the output files downloaded to a microcomputer using a communications packages. The resulting files can then be edited and inserted into the final report using a word processor.

General Guidelines for Comparing Statistical Packages

Selecting statistical packages is in large part a function of personal preference and experience. The key point to remember is to use the tool that best fits the given situation and with which the researcher feels most confident and comfortable.

There will be no attempt to provide an overview of all the major statistical packages that exist. There are a variety of statistical packages, some specialized and some with only regional popularity and support that simply cannot be mentioned here due to limited space. The following points, however, will be noted.

Data Management Needs: How much data manipulation, auditing, and preliminary analysis need to be performed? The importance of this question cannot be stressed too much. It is these functions that can take up as much as 90 percent of the researcher's time and in some instances make or break a study.

Data Interchange Capabilities: Can the statistical package read and write data stored in a variety of formats? How easy would it be to transfer data from one package to another? For example, SAS and SPSS-X can read each other's data sets. At minimum, a statistical package should have the capability of reading and writing an ASCII file. Further, some packages such as SAS can read data in a variety of formats, such as packed fields, etc.

Frequency of Use: How often will the package be used? Will it be used for only one analysis or for a series of analyses? Will it be used on a day-in, day-out basis or only infrequently? If a package is to be used only infrequently, then ease of use becomes an important factor.

Ease of Use: How easy is the package for the experienced user? Can descriptive statistics, such as means standard deviations, frequency counts, and cross tabulations be easily generated? Are the basic procedures for managing the data easy to use? This, however, can be taken too far. For example, is the package so easy to use that a novice user can perform complex sophisticated analyses without a basic understanding of the procedures involved?

Once the analyses have been performed, how easy is the output to read and interpret? Is the output produced in a format that can be easily summarized or converted into a final report format?

Documentation and Support: Is the documentation available? Are the algorithms that are used adequately documented? Is the documentation divided by skill level? Putting information needed by the novice and experienced user in separate sections or manuals can be extremely helpful. Is local support and consulting available within the institution? Is external support available? Is there a national hotline or consulting service? Is training available? Is the package regularly and consistently maintained? Are new releases of the package being regularly installed and are they upward compatible (will the old commands still work with the new versions)? How widely is the package used locally? This last point should be considered with the following reservations. A number of packages are widely used simply because they have been around for so long, not necessarily because they are the best package for the job. A researcher should not be afraid to use a less well-known package if he feels strongly that it meets a particular need. In this game, variety can be used to great advantage.

Available Computer Resources: What quantities and types of computer resources are available? The best way to approach computer support is to select the software and then the hardware to run the selected software. This is unfortunately not what has been or is being done. Unless the researcher is in a position to purchase both hardware and software, he may be limited by available resources. Certain packages simply take more computer resources to

run than others. These differences also may not hold for the package as a whole, but may differ depending on the specific procedure that is being used. The disk space required to store data sets while they are being analyzed is often also a very important consideration.

Efficiency and Cost: What are the total costs of using the package? This is another point that cannot be considered too strongly. The total costs of using the package should always be considered: not only computer resources, but the salaries of the individuals who will be using the package and the time needed to accomplish a particular task.

The Range of Statistical Procedures: What is the range of statistical procedures contained in the package? If the particular procedure needed is contained in only one package, the decision may already have been made for the researcher. If all the procedures needed are contained in one package, then this can be a plus. If, however, the package has adequate data interchange capabilities, then a limited set of statistical procedures should not prevent the use of the package. It is not an uncommon practice to use a variety of packages when a series of analyses need to be performed.

MAINFRAME PACKAGES

Perhaps the three most widely used statistical packages on mainframe and minicomputers are BMDP, SAS, and SPSS. Each of these packages has its advantages and disadvantages. All three of these packages produce accurate and reliable results if they are used properly. These three packages will be briefly described in order of ease of use, from easiest to more difficult.

BMDP: BMDP Statistical Software, Los Angeles, CA. One of the original packages. Available on a wide variety of mainframes, minicomputers, and microcomputers. Limited data manipulation, data management capabilities. Documentation is adequate, but may prove a problem for the novice user. Reasonably easy to use once the syntax is mastered. Contains specialized procedures which simply may not be available on other packages. Limited external support.

SPSS-X: SPSS, Inc., Chicago, IL. Easy to use. Powerful. Limited, but useful data manipulation capabilities. Good data interchange capabilities. Clear, well written manuals, with examples. Manuals are available for a variety of user skill levels. Manuals are useful, not only for using the package, but can also serve as useful statistical primers. Available on a wide variety of mainframes. Wide variety of statistical procedures. Format of output can be a bit voluminous. Adequate, but limited graphics capabilities on the mainframe version. Exceptionally useful Hierarchical, Log-Linear, and Discriminant analysis routines. Limited external support. A good choice for novice

users or situations where limited data manipulation is required.

SAS: SAS Institute, Cary, NC. No statistical package currently on the market can match the data manipulation, data management capabilities of this package. This should be the package of choice if extensive data manipulation/data management is to be required. Extremely powerful merge and update capabilities. An extensive library of manuals designed for users of different levels. Basics and Introductory manuals are well written and easy to use. Statistics Guide assumes a level of sophistication on the part of the user that may or may not be present. A wide range of statistical procedures. Excellent graphics capabilities. Excellent data interchange capabilities; can read data files that other packages cannot. Harder to learn than SPSS-X and BMDP, but easy transition for old SPSS and BMDP users and well worth the extra time. Can require more machine resources than SPSS-X and BMDP, but usualy requires considerably less time for the experienced user to set up and use. Available only on a limited number of mainframes, minicomputers and microcomputers. Originally only available on IBM[4] or IBM look-a-like mainframes, but now available for VAX[6], Data General[8] and Prime[7] minicomputers. Good external support, including consulting, and training courses and materials. This should be the package of choice for the researcher who spends most of his time with data manipulation and data analysis.

MICROCOMPUTER PACKAGES

Microcomputer versions of the three major mainframe packages mentioned earlier do exist. The comments made with respect to the mainframe versions of these packages in general hold for the microcomputer versions. Minor, yet sometimes significant, changes have been made in the packages to allow their use on microcomputers. Some of these changes are worth noting and will be mentioned below. In addition, there is also one other package that deserves mention, Systat[15].

BMDPC: (Version 2.05) BMDP Statistical Software, Inc., Los Angeles, CA. A downsized version of the mainframe package. A close copy of the mainframe version. Minimally interactive. Not rewritten to take advantage of microcomputer's unique capabilities. No on-line help. Extremely limited data interchange capabilities, can read ASCII files, but not translate files from other packages. IBM PC or close compatibles, MS-DOS[5] 2.0 or later, requires 640K RAM, hard disk and math coprocessor. Useful for the experienced BMDP mainframe user, but not necessarily for the novice user. Copy-protected.

SPSS-PC+: An exceptionally good micro implementation of a mainframe package. The package is fast and reasonably easy to use with limited data manipulation capabilities, but can be a powerful package when coupled with a database

management package such as RBASE 5000[12]. Excellent data interchange capabilities including access to mainframe SPSS-X and SAS data sets, Lotus 1-2-3[9], Symphony[9], Dbase II and III[10], etc. Excellent documentation. External support appears to be considerably better for the PC than the mainframe version. Graphics, considerably better than the mainframe vesion. IBM PC's and close compatibles, PC-DOS[5] or MS-DOS 2.0 or later, 384K and a hard disk, math coprocessor recommended. A very good choice for the researcher who has limited access to mainframe packages and is dealing with data sets containing 5,000 observations or fewer. The package can be effectively run on an 8088 based machine, but for larger data sets, an 8086 or 80286 machine is recommended. Copy-Protected.

PC SAS: (Version 6.02) Minor syntax changes from mainframe Version 5. The same data management/manipulation capabilities as the mainframe version. On-line help. Windows. Limited program editor. Procedures can be run from menus. External support is limited, no direct external support for average user, must go through site representative. Slow, even on an 80286 based machine. IBM PC, AT or compatibles, PC-DOS or MS-DOS 2.0 or later, 512K RAM (640K recommended), one floppy and 10-megabyte hard disk (20-megabyte recommended), supports math coprocessor, but unlike SPSS-PC addition of coprocessor does not result in significant reduction in processing time. Documentation is massive, and while adequate for the experienced SAS user, may serve to scare off the novice user. For the experienced and devoted SAS user, this is unfortunately a disappointing first entry into the microcomputer arena. The only real advantage for PC SAS would be as a development and testing tool for mainframe programs. Not Copy-Protected, but license must be renewed each year, built in routine checks system clock.

SYSTAT: (Version 3.0) Systat Inc., Evanston, IL. A good package for exploring data. True interactive capabilities. Commands are rather terse, but ease of use increases as package is used. Highly accurate. Command files can be built using any ASCII editor. No graphics. Extremely good external support. Good data interchange capabilities. A good range of procedures. Limited data manipulation capabilities. Documentation is adequate, but could stand a few more examples. IBM PC, AT or compatibles, Apple Macintosh[16]. The PC version requires 256K RAM, PC-DOS or MS-DOS 2.0 or higher, two floppy disk drives, a hard disk is recommended, a math coprocessor is supported and can greatly decrease execution time. This is a fine package for the researcher who must do a lot of statistical analysis. Limited data management capabilities can often be overcome by coupling this package with a good microcomputer database management package. Not Copy-Protected

SPREADSHEETS

The spreadsheet is similar to a giant financial ledger sheet - it is ruled into rows and columns. These rows and

columns form cells in which a single piece of information can be stored - number, formula or label. The power of spreadsheets is in their ability to do formula calculations. These calculations may range from the simple adding of a column of numbers to complex calculations.

The point to remember in institutional research, however, is that the quality of a decision does not rest on the quantity of the data analyzed. Spreadsheets are a tool to be used, but there is no substitution for clear thinking and knowledge of the problem. A recent comparison of ten spreadsheets conducted by PC Magazine concluded that

> No Spreadsheet was so bad that smart minds couldn't overcome it. Likewise, no spreadsheet was so good that it would overcome bad thinking.

DOES THE OFFICE REQUIRE THE USE OF SPREADSHEETS?

If a spreadsheet is needed, the decision of which spreadsheet is best should be based upon the specific nature of the problems to be solved and the level of computer sophistication of the user. PC World (December 1985) categorized the world into two types of users: "those who simply want results - quickly and safely - and those who crave more powerful products so they can build ever more intricate applications." Another way to look at the use of spreadsheets is as follows.

Extensive day-in, day-out data manipulation.

While use of the computer and spreadsheets may vary from office to office, there are some offices that are heavily involved with large scale number crunching and intricate applications. Integrated packages such as Framework[10], Symphony and Enable[14], and the stand-alone packages SuperCalc3[18] and Lotus 1-2-3, give the power and flexibility needed to handle such applications. Built-in functions and "add-in programs," along with logistical aids allows the management of data - if the user only needs to concentrate on the problem and not the logistical problems of how it is to be handled, the work will go more smoothly.

Integrated software is basically composed of spreadsheet, database management, word processing, graphics, and asynchronous communications. The major types of integration are: multi-function packages, compatible series, operating environments, background utilities, and desk organizers. The advantage of multi-function packages is that there is considerable savings in buying one package, while the advantage of a compatible series is the need to buy only modules that are most likely to be used. The major facets of integrated software are: 1) data compatibility - enables one application to call on files created by another application, 2) concurrency - ability to pull up any function along with any other, and 3) command consistency.

Intermittent users.

Not all institutional researchers will need to use spreadsheets every day. Occasionally requests for data analysis do not require the use of integrated spreadsheet programs. In fact, one "complaint" about these programs is that they are too complex for the average user. Whether or not this is the case, there are a number of good "medium-range" spreadsheet programs that permit users with less complex data analysis needs to obtain results easily. Among these programs are Multiplan[5] and Lotus 1-2-3.

Just need an answer.

If you just need an answer and are not a computer programmer, a commercial spreadsheet template (pre-defined programs or sets of keystrokes to perform specific tasks or calculations written in the command language used by the spreadsheet) may be the answer. These are pre-written programs that prompt you for the information needed to answer specific questions. Many are designed to work with Lotus 1-2-3 or Symphony, and can perform the following tasks: accounting, project management, office management, budgeting and cash management, tax planning and preparation, and real estate analysis. Commercial templates are available for some of the traditional institutional research kinds of analyses, such as Enrollment Forecasting, Student Flow Models, Enrollment Planning, Faculty Workforce Analysis, and Financial Assessment and Planning (NCHEMS, Vantage Information Products, Inc.). Similar templates may also be available from individual researchers. Good sources for this type of information, for the user, are computer "clubs," computer bulletin boards, or professional organization special interest groups.

Expensive typewriters. At the bottom end of the scale of spreadsheet use, the computer is merely used as a typewriter - not always the most efficient use of an expensive piece of equipment. In this case, the user merely keystrokes in the data, then, at best, has the spreadsheet program add a column or two of numbers before printing out the table. Such a use makes more sense if the data can be downloaded from a mainframe and imported to the spreadsheet or if the data tables are extensive and must be updated and printed on a regular basis.

WHAT CAN A SPREADSHEET DO?

Spreadsheet programs, like calculators, have "memories" where information may be stored, along with functions and pre-programmed formulas. Different functions (simple sets of keystrokes) that are available on Lotus 1-2-3 and Symphony, for example, are adding a column of numbers, computing square roots, computing present net value, converting decimal values to integers, specifying the number of decimals to include, and (for Symphony) generating random numbers, which means that a Monte Carlo simulation can be done. Other functions available are

statistical functions including: determining minimum and maximum values, computing the average, counting the number of non-blank cells, and computing the standard deviation and variance; financial functions including: computing the present value of a constant stream of payments (ordinary annuity); computing the present value of an uneven stream of payments; computing payment if you know the present value, the interest rate, and the term; and computing the internal rate of return on an investment. Standard mathematical, trigonometric and logical functions are also often included.

A user can create his own functions by creating macros - a special kind of program that runs entirely within the spreadsheet.

GENERAL QUESTIONS FOR SPREADSHEET USE.

Price considerations - How much are you willing to spend? Is it worth it to spend extra money on functions you will use only occasionally?

How much memory and disk space do you have available?

How 'fast' does the program need to be?

POPULAR SPREADSHEET PROGRAMS

Multi-function Programs

Framework, Ashton-Tate, Culver City, CA. Requires 384K RAM, two disk drives. A complex program, worth it for day-in, day-out users. Fast and flexible. Relatively small number of spreadsheet cells (4,000 - 5,000). Effective tool for organizing a problem and breaking it down into its various components. Sparse matrix technology. IBM PC, AT and compatibles. New version is not Copy-Protected.

Symphony, Lotus Development Corporation, Cambridge, MA. Requires 320K RAM, two disk drives, one drive, and a hard disk. Very powerful. Sparse matrix technology. Window capabilities. Criticized by many reviewers for being difficult to use. IBM PC, AT, and compatibles. Copy-Protected.

Enable, The Software Group, Ballston Lake, NY. Requires 192K RAM, two disk drives, graphics adaptor, PC-DOS 2.0 or higher. Good speed and ease of use, originally some documentation problems. IBM PC, AT and compatibles.

JAZZ, Lotus Development Corporation, Cambridge, MA. The original integrated spreadsheet package for the Macintosh. Implementation leaves little room for data. Files limited to available memory. Windowing tricky to use. Reasoably powerful, but slow access. Data interchange works in only one direction, to JAZZ. Adequate printed documentation. No on-line documentation. Requires 512K RAM; two disk

drives or a hard disk, additional memory recommended. Apple Macintosh. Copy-Protected.

Excel, Microsoft Corporation, Bellevue, WA. Currently the most widely used integrated package for the Macintosh. A fast and efficient implementation. Larger spreadsheet than JAZZ. Compares favorably with Lotus 1-2-3. Basic features easy to use and learn. Excellent data interchange capabilities to and from Excel. Adequate manuals and on-line help. Quality graphics. There is currently no de-install for hard disk use. Requires 512K RAM, two disk drives or hard disk. Apple Macintosh. Copy-Protected; current copy-protection scheme is one of few weaknesses.

Stand Alone Spreadsheets:

VisiCalc, Software Arts, Inc., Wellesley, MA. Requires 192K RAM, one disk drive. The 'original' spreadsheet program. Lacks many of the built-in functions of newer programs. Apple II series, IBM PC, AT and compatibles.

SuperCalc3, Sorcim Corporation, San Jose, CA. Requires 96K RAM, two disk drives or hard disk. Very fast; very powerful; good graphics. Not Copy-Protected.

SuperCalc4, Computer Associates International, Inc., San Jose, CA. One of the most powerful spreadsheet packages on the market. A viable alternative to Lotus 1-2-3. Excellent graphics. Requires 256K RAM, MS-DOS 2.0 or higher, two disk drives. Supports a variety of graphics cards. IBM PC, AT, and compatibles. Not Copy-Protected.

Multiplan, Microsoft Corporation, Bellevue, WA. Requires 128K RAM, one disk drive. Inexpensive; good choice for the novice. Sparse matrix technology. Interfaces well with graphics package, Microsoft Chart.

Lotus 1-2-3, Lotus Development Corporation, Cambridge, MA. Requires 128K RAM, two disk drives or hard disk, color graphics card. Fast; many functions; good reputation; sparse matrix technology. IBM PC, AT and compatibles. Copy-Protected.

INTERFACING SPREADSHEETS AND STATISTICAL PACKAGES

Some problems may best be approached using both spreadsheets and statistical packages. This can take several forms.

For the microcomputer user, PC SAS and SPSS-PC come with data interchange modules that allow the packages to read and convert Lotus 1-2-3 files via the DIF[11] format. Using these modules, data can be easily exchanged. Data which are stored on microcomputer spreadsheets can be written out as ASCII files which are then uploaded to a mainframe to be used as input for mainframe statistical packages. The reverse process can also occur using such features as the IMPORT capabilities of Lotus 1-2-3 or Symphony.

Another use would be to save the analysis results from mainframe statistical packages, download these results, and use a spreadsheet to summarize and further analyze the results. For example, summary statistics from repeated fittings of multiple regression models may be downloaded for a closer examination and comparison of the R^2s, mean Square errors, or regression weights. The graphics capabilities of such packages as Lotus 1-2-3 and Symphony can often be very useful in such an examination.

Spreadsheets can also often be helpful in visually interpreting the results of analyses. For example, given a set of regression weights, it is quite easy to create a dummy set of data, use the regression weights to calculate expected values and plot the observed versus the expected values. Using a spreadsheet, the researcher can get almost immediate feedback. While the same thing can be accomplished with a mainframe statistical package, the researcher usually does not have the degree of control or quality of graphical display he would have with a microcomputer. Also, unless the mainframe package has interactive capabilities, it may take an hour or so for each job to run, if not longer.

A more mundane, but extremely useful, application of spreadsheets is for tabling the results from statistical analysis. For example, when repeated Log-linear models are being fitted, spreadsheets can be extremely useful for summarizing the results and calculating the changes in degrees of freedom and deviance. Even if the need is simply for tabling data, the use of a spreadsheet can often make this job easier. One useful approach is to use a spreadsheet to build summary tables which are then written as ASCII files that can be merged directly into a final report.

SUMMARY

Statistical packages and spreadsheets are basic tools for the institutional researcher. The researcher should realize that it is not always a question of whether to use a statistical package or a spreadsheet, but rather how each of these tools can be applied to get the job done. It is usually not a question of choosing one or the other, but rather of deciding which can best be applied at the various steps in the data analysis process. Such a decision depends not only on the technical capabilities of the packages, but perhaps more importantly on the abilities, experience, and understanding of the researcher.

While there is some overlap in the technical capabilities of statistical packages and spreadsheets, they were initially designed and intended for different uses. Spreadsheets were initially designed to help summarize and describe data. They are also particularly useful for "what if" analysis. Statistical packages, on the other hand, were designed to provide support for hypothesis testing

and to provide data summaries and descriptions beyond simple calculations or manipulations.

To fully satisfy all the steps in the data analysis process, the researcher is likely to need all the tools at his command, including both statistical packages and spreadsheets. It must be remembered that before any useful statistical procedures can be chosen, the researcher must first adequately explore and describe the data that is to be analyzed. It is at this step and in the summary and presentation steps that the use of spreadsheets in conjunction with statistical packages can most aid the researcher.

Perhaps the most important point to remember is that computer software should never be substituted for a clear understanding of the statistical and technical procedures and processes involved. The easiest software package to use is not always the most appropriate one for the job at hand.

REFERENCE SOURCES

Computer software is, to borrow a phrase, perishable. The packages are continually changing and being updated. Thus books and articles reviewing such packages are usually outdated shortly after they hit the market. Publications that provide reviews of software on a regular basis include the following:

Review Sources for Mainframe Packages:

Datamation, Datamation, 875 Third Ave., New York, NY, 10022. The Systems Software review issue usually appears in December.

Data Sources, Ziff-Davis Publishing Co., One Park Avenue, New York, NY, 10016.

Datapro Reports on Software, Datapro Research Corporation, Delran, NJ, 08075.

Review Sources for Microcomputer Packages:

Byte, McGraw-Hill, Inc., P.O. Box 590, Martinsville, NJ, 08836.

PC World, PC World Communications, Inc., P.O. Box 55029, Boulder, CO, 80323-5029.

Infoworld, Infoworld, P.O. Box 1018, Southeastern, PA, 19398.

For a list of other sources, including a description of on-line indexes and services the following is recommended:

Glossbrenner, A. How to buy software. New York: St. Martin's Press, 1984.

Manuals for Mainframe Statistical Packages:

There are a multitude of manuals for the major statistical packages. These manuals tend to change as the versions of the packages change. Rather than attempt to list the various manuals, a source for the manuals or information on the manuals will be given.

BMDP - BMDP Statistical Software, 1964 Westwood Blvd., Suite 202, Los Angles, CA, 90025.

SAS - SAS Institute Inc., Publications Sales Department, Box 8000, SAS Circle, Cary, NC, 27511-8000.

SPSSX - SPSS Inc., 444 N. Michigan Ave., Chicago, IL, 60611.

Selected Reviews:

Fridlund, A. J., Statistics Software, Special Report, Infoworld, (1986), 8(35), pp. 31-39.

BMDPC (Version 2.05), BMDP Statistical Software, Inc., 1964 Westwood Blvd., Suite 202, Los Angeles, CA, 90025. Review: The American Statistician, (1985), 39(3), pp. 213-215.

MSUSTAT, Research and Development, Inc., Montana State University, Bozeman, MT, 59717-0002. Review: The American Statistician, (1985), 39(1), pp. 72- 74.

NCCS, Number Cruncher Statistical Systems, 865 East 400 North, Kaysville, UT, 84037. Review: The American Statistician, (1985), 39(4), pp. 315-318.

SYSTAT, Systat, Inc., 603 Main St., Evanston, IL, 60202. Review: The American Statistician, (1985), 39(1), pp. 67-70.

[1]SAS is a registered trademark of SAS Institute;

[2]SPSS-PC and SPSS-X are registered trademarks of SPSS Inc.;

[3]BMDP is a registered trademark of BMDP Statistical Software;

[4]IBM, IBM PC, and IBM AT are registered trademarks of International Business Machines Corporation;

[5]MS-DOS and Multiplan are registered trademarks of Microsoft Corpoation;

[6]VAX is a registered trademark of Digital Equipment Corporation

[7]PRIME is a registered trademark of Prime Computer, Inc.;

[8]Data General is a registered trademark of Data General Corporation;

[9]Lotus 1-2-3 and Symphony are registered trademarks of Lotus Development Corporation;

[10]Framework, DBASE II and DBASE III are registered trademark of Ashton-Tate;

[11]DIF is a registered trademark of Software Arts Products Corporation;

[12]RBASE 5000 and RBASE System V are registered trademarks of Microrim, Inc.;

[13]SuperCalc3 is a registered trademark of Sorcim Corporation;

[14]ENABLE is a registered trademark of The Software Group;

[15]SYSTAT is a registered trademark of Systat, Inc.;

[16]Macintosh is a trademark of Apple Computer Corp.