DOCUMENT RESUME

ED 345 670 HE 025 582

AUTHOR Hale, Charles D.; Travis, Marvin T.
TITLE Academic Program Level Effects Assessment: A Model.
PUB DATE 18 Mar 92
NOTE 22p.; Paper presented at the National Conference on Successful College Teaching and Administration (Orlando, FL, March 1992).
PUB TYPE Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Accountability; Admission (School); Alumni; *College Outcomes Assessment; Demonstration Programs; Departments; Educational Assessment; Higher Education; Program Descriptions

ABSTRACT

In response to increasing demands for evidence of institutional effectiveness in higher education, and given the diversity within the Academy, a decentralized model of institutional academic effects assessment, the Academic Program Level Effects Assessment Model (APLEAM), may be useful. APLEAM consists of three interdependent components: instructional, matriculation and post-graduation effects assessment. The instructional effects assessment component is carried out using criterion referenced testing to assess program and course level instructional effects because this approach can evaluate student mastery of knowledge, attitudes or skills taught; because the essential infrastructure for a criterion referenced assessment system exists in most institutions; because it provides an absolute statement of student attainment; and because it allows each institution to configure its assessment system to its unique curriculum. The student matriculation component of APLEAM includes program entry, program persistence, and program exit. APLEAM addresses post-graduation effects assessment through an alumni survey instrument which asks alumni to rate the program's personal and professional effectiveness, to identify ways that the major has been most and least helpful, to suggest program modifications, to indicate useful graduate competencies, and to identify graduate school intent or enrollment. (23 references) (JB)

Academic Program Level Effects Assessment:

A Model

By:

Charles D. Hale, Ed.D., C.H.E.S.
Assistant Professor
Healthcare Administration

and

Marvin T. Travis, D.P.A.
Professor of Management

Division of Business Administration
Saint Leo College
St. Leo, FL 33574

18 March 1992

Abstract

Academic Program Level Effects Assessment: A Model

There have been increasing demands for evidence of institutional effectiveness. Given the diversity within the Academy, a "top/down" monolithic assessment model of academic effectiveness is of limited value. A decentralized model of institutional academic effects (outcomes) assessment is presented for consideration. The Academic Program Level Effects Assessment Model (APLEAM) consists of three interdependent, components: instructional, matriculation, and post-graduation effects assessment. Model assumptions are presented and examined. Each component of APLEAM is described and discussed within the context of prevailing practice and theory.

Introduction

An assessment of institutional academic effectiveness is mandated by regional accreditation agencies such as the Southern Association of Colleges and Schools (1987a, p. 11; 1987b, p. 10). As higher education institutions are characterized by great diversity within their components, faculties, students, and other publics, the assessment of institutional effectiveness is not a simple task.

In assessing institutional effectiveness, a top-down, monolithic assessment model would be of marginal utility as such an approach tends to obscure the effects of unique academic program characteristics and/or other potentially significant, but modest effects.

It is the central thesis of this paper that an assessment of institutional academic effectiveness is best engaged at the academic program level (e.g., major). The basic assumptions underlying this approach are:

1. it is the faculty, based on their possession of the requisite special knowledge and proximity to students, who are best equipped to assess the effects of academic programs;

2. the curriculum and faculty exit separate and apart from delivery systems, however configured. Hence, an academic effects assessment will cut across curricular delivery system lines in institution's with multiple delivery systems; and

3. it is necessary to "start where the institution is" with respect to the design, implementation, and utilization of any effectiveness assessment activity (support at the presidential level is absolutely essential).

An Approach for Assessing Institutional Effectiveness

An institutional effectiveness assessment approach is

profiled in Figure 1, On the vertical axis, are four hierarchical levels. The lowest level (program), may be defined to be an academic department, major, minor, or concentration. The next level up (intermediate), may be defined to include a division, school, or college depending on the institution involved. The third level (executive), refers to an academic vice-president, or equivalent, who is interested in assessing only academic and/or academic support services. The highest level is that of the institution's chief executive officer (CEO), i.e., the president who reviews effectiveness assessment reports from all the vice-presidents (or equivalent) of the institution. At the lowest levels of the model, effectiveness assessment studies are narrow in scope. As one moves vertically up the model, effectiveness assessment studies expand in scope, until at the CEO level, the whole institution is considered.

For example, within an liberal arts college's division of business administration, there exist three majors: business, healthcare, and human resources administration. Given the outline presented in Figure 1, the faculty teaching within the healthcare administration major, as well as those teaching within the other two majors, would conduct three independent program effect assessment studies (program level). The results of these three studies would be summarized and combined into a single report by the division chair (intermediate level). Since the liberal arts college has eight academic divisions, eight separate divisional

reports would be submitted to the office of the Vice President for Academic Affairs (executive level). The Vice President would summarize and combine the divisional reports into a single institutional academic effects assessment. The other vice-presidents (e.g., student affairs and operations) would follow a similar process for their areas of responsibility. Finally, the vice-presidents would submit their reports to the President (CEO level) who would combine the reports into a single institutional effectiveness assessment report.

INSERT FIGURE ONE The Academic Program Level Effects Assessment Model

The purpose of this paper is to describe a model for assessing program level effects upon students. For the purposes of the Academic Program Level Effects Assessment Model (APLEAM), (Figure 2), program level may be defined to include either a major, minor, cognate, concentration, or specialization. The model is segmented into three assessment components: instructional, matriculation, and post-graduation effects.

INSERT FIGURE TWO

The assessment of instructional effects may be employed at either the course or program level. Within APLEAM, the assessment of instructional effects is conducted at both the course and program levels. Common instructional effects assessment techniques, at the course level, include unit and final examinations, presentations, and/or term papers. At

the program level (e.g., major), comprehensive written and/or oral examination(s), terminal projects, and/or capstone courses may be employed.

Assessing the effects of non-academic variables, during student matriculation, is desirable, as such data provides valuable information which assists in understanding the student's environment. Armed with such information, faculty are better able to design courses, select learning strategies, advise students, and interpret data obtained from the other two components of APLEAM.

Knowledge of post-graduation effects, i.e., assessment of the program's utility given personal and professional experience gained after graduation, are very important and can be instrumental in improving the quality of an academic program.

The potential uses of a program level effects assessment model include: (a) improved program management capacity, (b) improved effectiveness in student recruitment and retention, and (c) enhanced academic quality.

Instructional Effects Assessment Component

Introduction

Academic programs, within higher education institutions, are predicated upon the assumption that there exist domains of knowledge, attitudes, and/or skills for students to master. These domains of knowledge, attitudes, and/or skills are segmented into courses which are then bundled into majors, minors, specializations, cognates, or concentrations.

There are essentially two approaches to the assessment of student mastery over specified domains: norm and\or criterion referenced testing. Such testing may be administered via written examination and/or application, e.g., thesis, project, or skills demonstration.

Testing: An Introduction

A norm has been defined as, "a single value, or distribution of values, constituting the typical performance of a given group" (The Joint Committee on Standards for Educational Evaluation, 1981, p. 154). Norm referenced testing requires the comparison of a student against the performance of other students across a specified domain of interest in relative terms, i.e., as the nature of the normed group changes so does the performance level of an examinee with respect to the reference group (Crocker and Algina, 1986, p. 69).

The second approach to assessing student mastery, over defined domains of knowledge, attitudes, and/or skills has been commonly referred to as "criterion referenced testing" Popham (1981, p. 30). Instructional objectives are used to define the domain of interest (Osterlind, 1989; Oosterhof, 1990). A specified, absolute performance standard is imposed which is unrelated to a normed group's performance. As long as the content of the specified domain remains stable, a student's mastery score, (e.g., percentage of test items correct) is an absolute reflection of his or her degree of domain mastery.

The determination of cut off scores (also called standards) is externally imposed. There are a variety of mechanisms for standard setting, which range from expert judgment to complex statistical models (Crocker & Algina, 1986, pp. 411-428.) Almost all institutions of higher education have in place grading systems which specify levels of student mastery over defined bodies of knowledge, attitudes, and/or skills. Levels of student mastery have been expressed as a letter (A, B, C, D, or F), weighted average (e.g., 93 or 56), or as either pass/fail.

Within APLEAM, a criterion referenced testing approach is advocated to assess program and course level instructional effects because: (a) a purpose of "outcomes assessment" is to determine student mastery of the knowledge, attitudes, and skills taught, (b) the essential infrastructure for a criterion referenced assessment system exists in most, if not all, higher education institutions, (c) the ability to make absolute statements of student attainment at a particular institution is afforded, and (d) it seems appropriate for each institution to configure its assessment system so as to be positioned to assess its unique curriculum.

If a combined (norm and criterion referenced) approach is employed, a normed examination can be administered so as to learn how students rank. A criterion referenced achievement test can be given so as to identify how students perform relative to an institution's standards and as well as for diagnostic purposes. In the final analysis, the approach

(or combination) that is selected should be dependent upon what questions need to be answered.

Testing: Technical Considerations

Content Validity

Criterion referenced testing, like norm referenced measurement, is concerned about content validity which has been defined as, "validation for situations where the test user desires to draw an inference from the examinee's test score to a larger domain of items similar to those on the test itself," e.g., level of an examinee's domain mastery (Crocker and Algina, 1986, 217). A test's content validity is dependent upon the degree to which its items are reflective of, i.e., drawn from, its corresponding content domain and that domain's content homogeneity (Crocker and Algina (1986, p. 218).

For example, suppose an achievement test was designed to assess student mastery of basic accounting content in a course. Further, suppose that a content validation study was conducted and established the test's content validity. However, over time, basic accounting content in the course was replaced with marketing content. The test remained unchanged. Of course, the level of student mastery, as a function of the test, would probably decline as the domain's content homogeneity had been violated. Basic accounting content had been replaced with marketing content.

Steps must be taken to ensure a test's content validity and the corresponding domain's content homogeneity. This is

most often accomplished through the specification of explicit test development guidelines and instructional objectives (Crocker and Algina, 1986, pp. 66-83 & 218). Other, techniques and/or processes to ensure content homogeneity include: (a) course curriculum guides (prepared by content experts) with minimum competencies subsumed under each specified learning objective, (b) faculty screening procedures, (c) peer review of syllabi, (d) transfer credit equivalence evaluation, (e) course equivalence and waiver determinations, and (f) technical assistance with assessment and evaluation issues.

Reliability Estimation

Criterion referenced tests must possess reliability. However, there is disagreement as to whether or not criterion referenced tests may employ reliability estimation techniques originally developed for norm referenced tests (Popham and Husek, 1969; Oosterhof, 1990, p. 209).

A norm referenced estimation reliability technique is suggested (at least in initial test development) for the following reasons: (a) the utility of reliability theory mastery classification techniques is limited as most institutional grading systems have an explicit cut score that determines mastery level; (b) the interpretation of test scores [in criterion referenced testing] does not require comparison, against other examinees, as test scores can be interpreted even if all examinees earned the same score (Oosterhof, 1990, p. 209); and (c) the apparent limited

state-of-the-art of criterion referenced testing indices as described by Subkoviak, 1982, pp. 129-185).

Cronbach's alpha can be used to estimate internal consistency (reliability) based on a single administration (Crocker and Algina, 1986, pp. 138-139). Once, the desired examination has been fully developed, mastery classification theory (a criterion referenced reliability index) may be applied (Crocker & Algina, 1986, pp. 197-203).

Evaluating Test Item Function

Bloom's taxonomy has been used for years to guide item writing. The language of the taxonomy makes it very difficult to write items which assess student mastery beyond the its lower levels (Osterlind, 1989, p. 92). Seddon (1978) and Blumberg, Alschuler and Rezmovic (1982) have cautioned against the use of Bloom's Taxonomy until the levels within the taxonomy have been firmly established. Osterlind (1989, p. 93) has expressed agreement. Accordingly, APLEAM suggests the Framework for Instructional Objectives developed by Hannah and Michaelis (1977), as presented in Osterlind (1989, pp. 134-143).

It is essential that the appropriate test item format be employed depending on the knowledge, skill, or attitude being assessed. If the purpose of the examination is to assess student knowledge (i.e., recall of content), then a multiple choice, true/false, short answer, or other similar item format may be utilized. If the purpose of the test is to assess student attitudes, then a Likert scale, rating scale,

or a semantic differential may be employed. If the purpose of the examination is to assess student skill performance, then a writing sample, thesis, application project, simulation, case, exhibition, or problem sheet may be employed.

Test items, scored as either correct or incorrect, maybe assessed through application of classical item analysis techniques [item difficulty, distractor analysis, and item discriminating ability] (Crocker and Algina, 1986, pp. 329-330) or item response theory (pp. 339-371). APLEAM advocates: (a) the criteria for constructing test items advanced by Osterlind (1989) and (b) the test item editorial guidelines as outlined by Oosterhof (1990).

Student Matriculation Assessment Component

Introduction

Matriculation is defined as, "to enroll, especially as a student or candidate for a degree in a college." (Webster's New World Dictionary, 1966, p. 907). The theoretical basis for this APLEAM component is based, largely, on Cross' Chain of Response (COR) Model (1981, p. 124). The COR Model's utility has been demonstrated by Cooke (1986), Goodman (1983), Kraskouskas (1986), Hale, Nutter, Zemlo, and Willett (1989) Hale and Wattenbarger (1990), and Willett (1984).

When student matriculation behavior is examined, one should consider factors such as student (a) perceptions of academic ability, (b) attitudes towards education, (c) perceived goals and expectation of meeting those goals, (d)

the influence, if any, of life transitions on matriculation, (e) perceived opportunities and barriers, and (f) sources of information about matriculation opportunities and barrier removal (Cross, 1981, p. 124). Other factors to be considered would be education needs [depends on profession and/or worksite], previous educational attainment and age (Anderson & Darkenwald, 1979) as well as perceptions of academic program quality.

Matriculation Phases

Given these considerations, the student matriculation component of APLEAM includes: (a) program entry, (b) program persistence, and (c) program exit.

Phase 1: Program Entry

Students entering the program have varied demographic characteristics, academic goals, academic skill levels, and other concerns. Further, differing institutional and academic program features either serve to attract or repel prospective or newly enrolled students. Studies as to why new students enrolled can be designed and conducted for relatively little cost. Thus, such studies are within the reach of most higher education institutions.

The purpose of assessment within the program entry phase is to identify factors which induced new students to enroll and assess the effects of institutional and academic program characteristics on student enrollment decisions.

Phase 2: Program Persistence

Once students have enrolled and are matriculating within

the program, the matriculation issue becomes persistence, i.e., to what extent do students remain enrolled to accomplish educational objectives and/or graduate. The use a college's management information system (MIS) coupled with supplemental studies, designed to answer specific questions, should provide information as to the character and duration of student persistence.

Phase 3: Program Exit

During this stage, students are preparing to exit the program, with a substantial amount of matriculation experience, and are in position (a) to render judgments concerning the program; (b) to describe the influence of institutional and academic program characteristics on matriculation; (c) to describe the influence of selected variables on retention decisions; and (d) to report perceptions of academic quality.

Mechanism for Collecting Data

An efficient mechanism for collecting assessment data is an anonymous survey instrument which is completed by each student enrolled in the academic program at specifically designated points of matriculation.

The questions contained within each survey instrument may be subdivided into three categories (a) demographic items, (b) experiential items which are designed to be compared over time, and (c) items which have only been experienced by the group being studied.

Demographic items are needed so that it is possible to describe each group responding to a particular survey. Experiences, common to each group entering or leaving a program phase can be compared, thus providing data for use in assessing the influence of various factors affecting matriculation behavior. Questions which elicit descriptions of unique group experiences provide insight into their thoughts and behaviors. Following are descriptions of three possible survey instruments.

Student Profile (Phase 1). This instrument should produce data which describes the entering students age, gender, household membership, previous academic matriculation, occupational characteristics, travel patterns, residence, the influence of various institutional variables on the enrollment decision, and the subject's perception of successfully completing relevant academic exercises. Data are used to build a profile of entering students and assess the influence of institutional variables on student enrollment decisions. Implications for recruiting may emerge.

Student Flow Analysis (Phase 2). A student flow analysis will consider such factors as matriculation duration, credits consumed, "stop-out," and "drop-out" behaviors as suggested by Willett (1982). There can exist, on the student profile and/or senior exit interview instruments, questions designed to collect data which are used as indicators of student "flow" through an institution. Student flow data can be used to identify where within the

matriculation experience the program can intervene so as to ensure that students persist.

Exiting Student Survey (Phase 3). The purpose this survey is to ascertain exiting students' judgments of (a) the program's utility and quality, (b) access to and quality of student support services, and (c) specific factors affecting students' matriculation behavior. The survey instrument should contain both objective and subjective questions. The responses to subjective questions should be aggregated by question and subjected to content analysis. The survey instrument can gather data for evaluative purposes and can be administered in a capstone course.

Data Analysis Plan

Each survey instrument should be reviewed prior to data entry. A code book should be produced, as responses to survey items can be reduced to numeric codes so as to facilitate data processing. Level 1 data analysis consists of appropriate descriptive statistics or cross-tabulations, given the item. Level 2 data analysis consists of custom analysis using appropriate inferential (e.g., t-tests or correlations) or nonparametric statistic(s) to answer questions raised during Level 1 analysis.

Post-Graduation Outcomes Assessment Component

It is considered important to ascertain how well matriculation through the program's course work has prepared alumni for their professional and/or graduate education pursuits. Those who are best qualified to render such

appraisals are program alumni.

Alumni Profile. Questions on the survey instrument which operationalizes this component of APLEAM ask alumni to (a) rate the program's utility to them personally and professionally, given experience gained after graduation; (b) identify what within the major has proved most and least helpful; (c) suggest what content and skills should now be included or deleted within the program; (d) indicate what competencies graduating students should possess; (e) describe current occupation, job duties, and location; and (f) identify graduate school intent and/or enrollment. If graduate studies have been undertaken, data should be collected which describes alumni perceptions concerning the adequacy of preparation for graduate school.

References

Anderson, R. A. & Darkenwald, G. (1979). Participation and persistence in American adult education. New York: College Entrance Examination Board.

Bloom B. S. (1956). Taxonomy of educational objectives: The classification of educational goals. New York: Longman.

Blumberg, P., Alschuler, M. D., & Rezmovic, V. (1982). Should taxonomic levels be considered in developing examinations? Educational and Psychological Measurement, 42, 1.7.

Cooke, W. (1986). An application Cross' chain of response model describing faculty who participate in professional development activities (Doctoral Dissertation, University of Florida, 1986). Dissertation Abstracts International, 47/06A, 2041.

Crocker, L. & Algina J. (1986). Introduction to classical and modern test theory. New York: Holt, Rinehart, & Winston.

Cross, K. P. (1981). Adults as learners. San Francisco: Jossey-Bass.

Goodman D. L. (1983). Self-concept as a learner and level of threat in learning activities: Potential inhibitors of adults' participation in education (Doctoral Dissertation, Kansas State University, 1983). Dissertation Abstracts International, 44/12A 3573.

Hale, C. D. & Wattenbarger, J. L. (1990). Adult undergraduate academic participation. Community/Junior College Quarterly of Research and Practice, 14, (3), 199-211.

Hale, C. D., Nutter, R., Zemlo, J., & Willett, L. (1989). Examining adult participation in community school classes using Cross' chain of response model: Implications for recruiting. The Community Education Research Digest, 3, (2), 38-47.

Hannah, L. S., & Michaelis, J. U. (1977). A comprehensive framework for instructional objectives: A guide to systematic planning and evaluation. Reading, MA: Addison-Weesley.

Joint Committee on Standards for Educational Evaluation (1981). Standards for evaluations of educational programs, projects, and materials. New York: McGraw Hill Book Company.

Kraskouskas, G. (1986). Adult learning Practices: Transitions and stability (Doctoral Dissertation, Harvard University, 1986). Dissertation Abstracts International, 47/06A, 1982.

Oosterhof, A. (1990). Classroom applications of educational measurement. Columbus, OH: Merrill Publishing.

Osterlind, S. J. (1989). Constructing test items. Boston: Kluwer Academic Publishers.

Popham, W. J. & Husek, T. R. (1969). Implications of criterion referenced measurement. JEM, 6, 1-9.

Popham, W. J., (1981). Modern Educational Measurement. Englewood Cliffs, NJ: Prentice-Hall.

Seddon, G. M. (1978). The properties of Bloom's taxonomy of educational objectives for the cognitive domain. Review of Educational Research, 48, 303-323.

Southern Association of Colleges and Schools. (1987a) Criteria for Accreditation. Atlanta: Author.

Southern Association of Colleges and Schools. (1987b) Resource Manual on Institutional Effectiveness. Atlanta: Author.

Subkoviak, M. J. (1982). Decision consistency approaches. (In R. A. Berk ed. Criterion Referenced Measurement: The State of the Art). Baltimore: The Johns Hopkins University Press.

Webster's New World Dictionary (1966). New York: Merriam.

Willett, L. H. (1982). Continuing education student flow analysis. Research in Higher Education, 17, (2), 155-164.

Willett, L. H. (1984). Analysis of external degree student participation and attitudinal variables. Innovative Higher Education, 9, (1) 48-58.

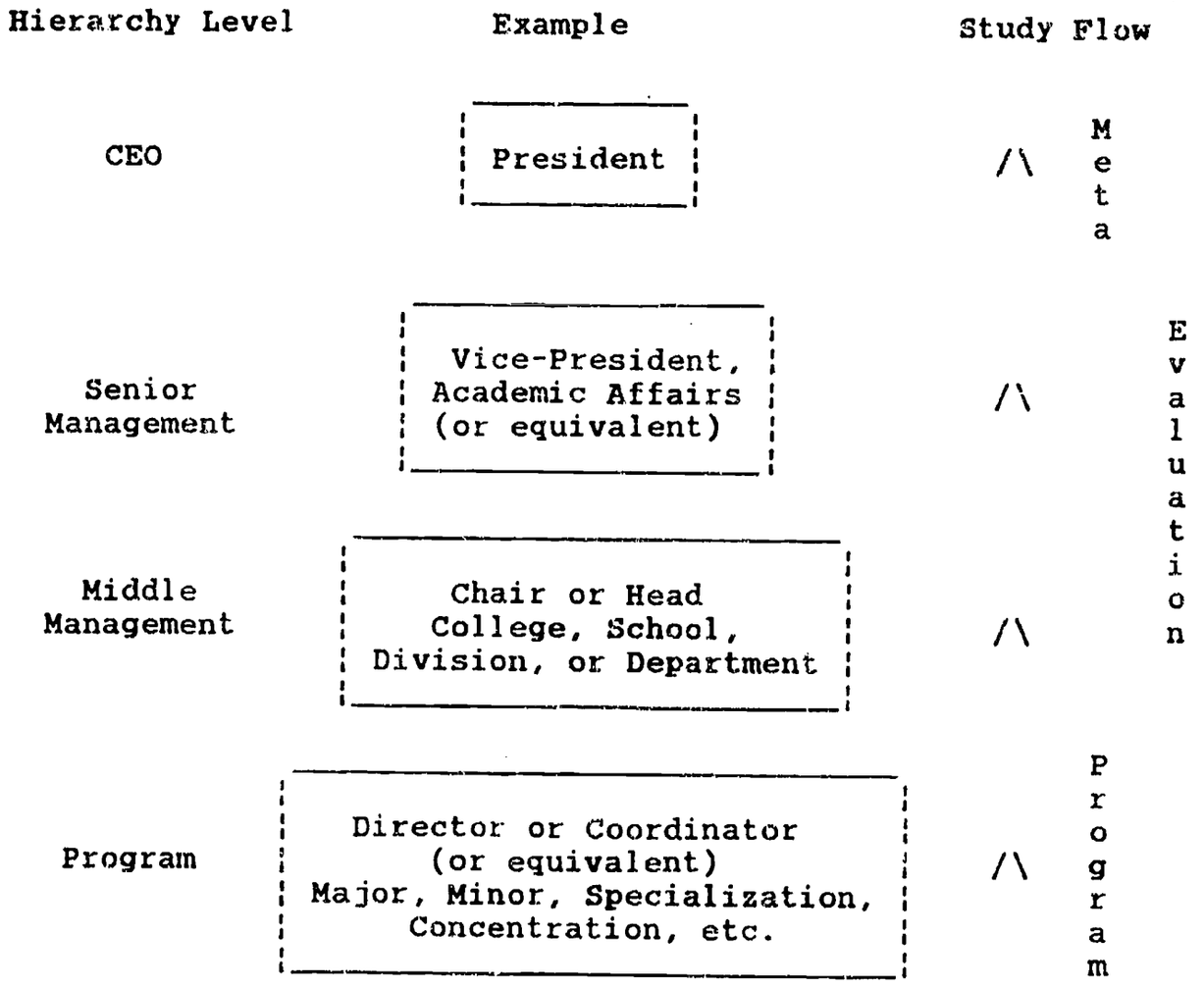

Figure 1. An Institutional Effectiveness Assessment Approach

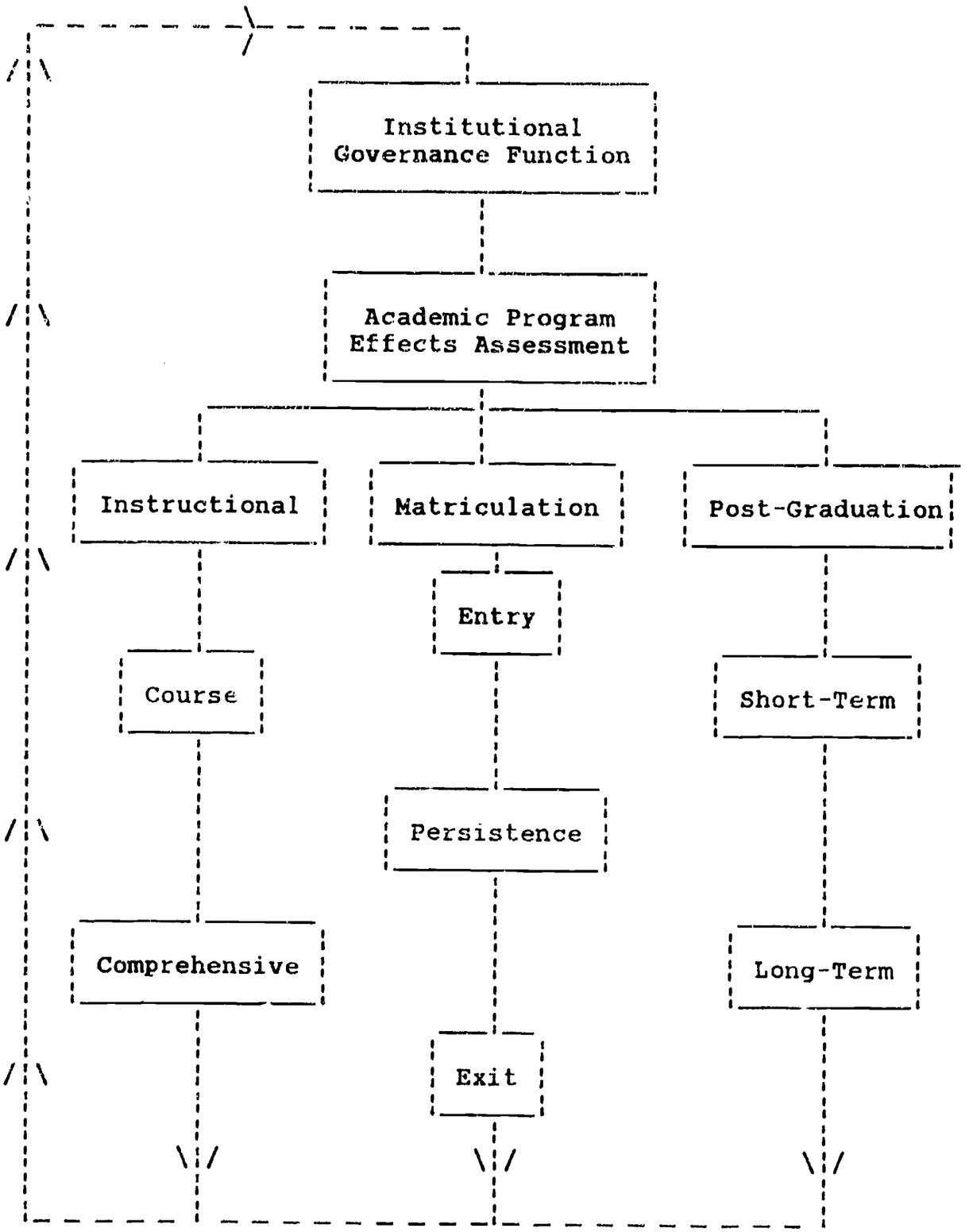

Figure 2. Academic Program Level Effects Assessment Model